PRESENTATIONS OF FINITE SIMPLE GROUPS: A COMPUTATIONAL APPROACH

R. M. GURALNICK, W. M. KANTOR, M. KASSABOV, AND A. LUBOTZKY

Abstract. All nonabelian finite simple groups of rank n over a field of size q, with the possible exception of the Ree groups ${}^2G_2(3^{2e+1})$, have presentations with at most 80 relations and bit-length $O(\log n + \log q)$. Moreover, A_n and S_n have presentations with 3 generators, 7 relations and bit-length $O(\log n)$, while $\mathrm{SL}(n, q)$ has a presentation with 7 generators, 25 relations and bit-length $O(\log n + \log q)$.

Contents

2000 *Mathematics Subject Classification.* Primary 20D06, 20F05 Secondary 20J06.

The authors were partially supported by NSF grants DMS 0140578, DMS 0242983, DMS 0600244 and DMS 0354731. The research by the third author also was supported by Centennial Fellowship from the American Mathematics Society.

1. Introduction

In [GKKL1] we provided short presentations for all alternating groups, and all finite simple groups of Lie type other than the Ree groups ${}^2G_2(q)$, using at most 1000 generators and relations. In [GKKL2] we proved the existence of profinite presentations for the same groups using fewer than 20 relations. The goal of the present paper is similar: we will provide presentations for the same simple groups using 2 generators and at most 80 relations. These and other new presentations have the potential advantage that they are simpler than those in [GKKL1], at least in the sense of requiring fewer relations; we hope that both types of presentations will turn out to be useful in Computational Group Theory.

The fundamental difference between this paper and [GKKL1] is that here we achieve a smaller number of relations at the cost of relinquishing some control over the length of the presentations. Our first result does not deal with lengths at all:

Theorem A. *All nonabelian finite simple groups of Lie type, with the possible exception of the Ree groups ${}^2G_2(q)$, have presentations with 2 generators and at most 80 relations.*

All symmetric and alternating groups have presentations with 2 generators and 8 relations.

In fact, a similar result holds for *all* finite simple groups, except perhaps ${}^2G_2(q)$ (the sporadic groups are surveyed in [Soi]). Both the bounds of 20 relations in [GKKL2] and 80 here are not optimal – in all cases we will provide much better bounds, though usually with more generators. Possibly 4 is the correct upper bound for both standard and profinite presentations. Wilson [Wi] has even conjectured that 2 relations suffice for the universal covers of all finite simple groups.

Although we are giving up the requirement of length used in [GKKL1], we can still retain some control over a weaker notion of length used in [BGKLP, BCLO] and especially suited for Computer Science complexity considerations: *bit-length*. This is the total number of bits required to write the presentation, so that all exponents are encoded as binary strings, the sum of whose lengths enters into the bit-length. (The presentation length that had to be kept small in [GKKL1] involves at least the much larger sum of the actual exponents; cf. Section 2.)

Theorem B. *All nonabelian finite simple groups of rank n over a field of size q, with the possible exception of the Ree groups ${}^2G_2(q)$, have presentations with at most 14 generators, 78 relations and bit-length $O(\log n + \log q)$.*[1]

Here we view alternating (and symmetric) groups as having rank $n-1$ over "the field of size 1" [Ti1]. The bounds in both of the preceding theorems also hold for many of the almost simple groups.

[1] Logarithms will be to the base 2.

By [GKKL1, Lemma 2.1], if we have any presentation of a finite simple group G with at most 78 relations, we obtain a presentation with 2 generators and at most 80 relations (cf. Lemma 3.39 below). Moreover, the proof of that lemma shows that *any* pair of generators of G can be used for such a presentation (cf. Corollary 3.40). "Almost all" pairs of elements of a finite simple group generate the group [Di, KLu, LiSh]; some pairs will probably force the length or even the bit-length to be fairly large. Indeed, [GKKL1, Lemma 2.1] is so general that it allows us to cheat somewhat: the resulting presentations are not even slightly explicit, and we have no information concerning their bit-lengths. In particular, we are unable to prove Theorem B using 2 instead of 14 generators.

In view of [GKKL1, Lemma 2.1] or Lemma 3.39, our goal will be to prove Theorem B. Much better bounds are obtained in various cases. For example, Theorems 3.27 and 3.36 deal with the alternating and symmetric groups, where we go further than any previous types of presentations for these groups in terms of the small number of relations used (cf. [GKKL1, BCLO]):

Theorem C. *For each $n \ge 5$, A_n and S_n have presentations with* 3 *generators*, 7 *relations and bit-length $O(\log n)$, using a bounded number of exponents.*

Moreover, for the preceding groups, in addition to the second part of Theorem A we show that, *if a and b are* any *generators of $G = A_n$ or S_n, then there is a presentation of G using* 2 *generators that map onto a and b, with* 9 *relations* (Corollary 3.40). There are similar results in all cases of Theorem A (cf. Remark 4 in Section 11). However, as already noted, we do not know if it is possible to choose a and b in order to obtain a presentation with bit-length $O(\log n)$; nor if it is possible to choose them in order to obtain a bounded number of exponents for groups of Lie type over arbitrarily large degree extensions of a prime field (cf. Remark 8 in Section 11).

In order to obtain all of the presentations in the preceding theorems, although [GKKL1] was a starting point we need significantly new methods for unbounded rank n; these ideas may prove to be more practical for actual group computation than some of those in [GKKL1]. Moreover, while a few of the arguments used here are streamlined, often simpler, and occasionally improved versions of ones in [GKKL1], they are still rather involved. As in [GKKL1] we do not use the classification of the finite simple groups in any proofs.

For groups of bounded rank, our presentations can be made to be short in the sense used in [GKKL1], at the cost of adding a small number of additional generators and relations (so that [GKKL1, (3.3) and (4.16)] will apply; cf. (3.15) and (4.3) below). It is our treatment of unbounded rank that contains new ideas to decrease the number of relations in [GKKL1] at the expense of the length of the presentation. We provide more than one approach for this purpose: some classical groups are handled in different ways in Sections 9 and 10. The unitary groups are dealt with separately in Section 8 by using an idea of Phan [Ph] as improved in [BeS].

In Sections 3–10 we will consider various types of simple groups in order to prove the above theorems, providing better bounds for the number of generators and relations in various cases. For many cases we only give hints regarding the final assertion in Theorem B.

One of our original motivations for work on presentations was the following

Corollary D (Holt's Conjecture [Ho] for simple groups)**.** *There is a constant C such that, for every finite simple group G, every prime p and every irreducible $\mathbb{F}_p[G]$-module M, $\dim H^2(G, M) \le C \dim M$.*

This conjecture has already been proven twice, in [GKKL1, Theorem B′] and [GKKL2, Theorem B]. As in [GKKL1] it is an immediate consequence of Theorem A, except for the Ree groups (and these also had to be handled separately in [GKKL1]). The proof based on the present paper (using the elementary result [GKKL1, Lemma 7.1]) is simpler than the previous proofs, although a smaller, explicit constant C is given in [GKKL2]. See [GKKL2, Theorem C] for a generalization of the preceding result to all finite groups.

Section 11 contains further remarks concerning these results. For now we note one further unexpected direction:

Efficient presentations. If $\langle X \mid R\rangle$ is a presentation of a finite group G, then $|R| - |X|$ is at least the smallest number $d(M)$ of generators of the Schur multiplier M of G; and $\langle X \mid R\rangle$ is called an *efficient* presentation if $|R| - |X| = d(M)$ [CRKMW, CHRR1, CHRR2, CHR]. The only infinite families of nonsolvable groups known to have efficient presentations appear to be groups having $\mathrm{PSL}(2, p)$ as a composition factor when p is prime [Sun, CR3] (cf. (3.17). Therefore it may be of some interest that Corollaries 3.8(i) and 3.12(ii) contain examples of families of groups having efficient presentations with alternating groups as composition factors. For example, *for any prime $p \equiv 11 \pmod{12}$, there is a a presentation of $A_{p+2} \times T$ with 2 generators and 3 relations, where T is the subgroup of index 2 in $\mathrm{AGL}(1, p)$.* It seems plausible that this can be used to obtain efficient presentations of A_{p+2} with 3 generators and 4 relations for these primes.

Examples 3.16 and 3.19, together with Table 1 and Remark 4.8, deal with presentations for various groups A_n and S_n when n has a special form. Section 3.5 contains explicit presentations of S_n for all $n \ge 50$. For general n it would be desirable to have even fewer relations than in Theorem C, with the goal of approaching efficiency for alternating groups.

2. Preliminaries

Presentation lengths. In [GKKL1, Section 1.2] there is a long discussion of various notions of "lengths" of a presentation $\langle X \mid R\rangle$ and some of the relationships among them. Here we only summarize what is needed for the present purposes.

length = word length: $|X|+$ sum of the lengths of the words in R within the free group on X. Thus, length refers to strings in the alphabet $X \cup X^{-1}$. This is the notion of length used in [GKKL1], and seems the most natural notion from a purely mathematical point of view. We reserve the term *short presentation* for one having small length. Achieving this was one of the goals in [GKKL1], though not of the present paper.

bit-length: the total number of bits required to write the presentation, used in [BGKLP] and [BCLO]. All exponents are encoded as binary strings, the sum of whose lengths enters into the bit-length, as does the space required to enumerate the list of generators and relations.

expo-length: the total number of exponents used in the presentation.

By comparing the present paper with [GKKL1] it becomes clear that small bit-length is much easier to achieve than small length. The properties required of the bit-length $bl(w)$ of a word w are

$$bl(x) = 1, x \in X; \quad bl(w^n) \le bl(w) + \log |n| \ \text{ if } \ n \in \mathbb{Z} \setminus \{0, 1, -1\}\,; \ \text{ and}$$
$$bl(ww') \le bl(w) + bl(w')$$

for any words w, w'.

Subgroups. We will use the elementary fact [GKKL1, Lemma 2.3] that a group G that has a presentation based on a known group, using presentations of subgroups of that group, has the subgroups automatically built into G:

Lemma 2.1. *Let* $\pi\colon F_{X\cup Y} \to G = \langle X, Y \mid R, S\rangle$ *and* $\lambda\colon F_X \to H = \langle X \mid R\rangle$ *be the natural surjections, where* H *is finite. Assume that* $\alpha\colon G \to G_0$ *is a homomorphism such that* $\alpha\langle \pi(X)\rangle \cong H$. *Then* $\langle \pi(X)\rangle \cong H$.

In the present paper we will use this freely, often without comments.

Curtis-Steinberg-Tits presentation. This is a standard presentation for groups of Lie type; see [Cur], [St1], [Ti2, Theorem 13.32] and [GLS, Theorem 2.9.3]. We will generally refer to [GKKL1, Sections 5.1 and 5.2] for a discussion of the versions we will use.

3. Symmetric and alternating groups

We will use a presentation for alternating groups, due to Carmichael [Carm, p. 169], that is more symmetrical than a presentation due to Burnside and Miller ([Bur, p. 464], [Mil, p. 366]) in 1911 that was used in [GKKL1, (2.6)]. Moreover, Carmichael's presentation requires less data (i.e., fewer relations):

$$A_{n+2} = \langle x_1, \dots, x_n \mid x_i^3 = (x_i x_j)^2 = 1 \ \text{ whenever } \ i \ne j\rangle, \tag{3.1}$$

based on the 3-cycles $(i, n+1, n+2)$, $1 \le i \le n$. We will also use the standard presentations

$$A_4 = \langle x, y \mid x^3 = y^2 = (xy)^3 = 1\rangle \ \text{ and } \ A_5 = \langle x, y \mid x^5 = y^2 = (xy)^3 = 1\rangle \tag{3.2}$$

[CoMo, p. 137]. Presentations are known for the universal central extension of A_n, $4 \le n \le 9$, with 2 generators and 2 relations [CHRR1, CHRR2]; and for A_{10} using 2 generators and 3 relations [Hav] (cf. Example 3.19(10)). These can be used in some of our presentations in Sections 6 and 8 in order to save several relations.

In Section 3.1 we make crucial use of the symmetry of (3.1), as follows. Let $T = \langle X \mid R\rangle$ be a group acting transitively on $\{1, \dots, n\}$, viewed as acting on $\{1, \dots, n, n+1, n+2\}$. Introduce a new generator z corresponding to the 3-cycle $(1, n+1, n+2)$. We use additional relations in order to guarantee that $|z^T| = n$ in our presented group, as well as a very small number of relations of the form $(zz^t)^2 = 1$ for suitable $t \in T$, in order to use (3.1).

This idea is reworked several times in order to handle various special degrees n. We glue two such presentations in Section 3.4 in order to deal with symmetric and alternating groups of arbitrary degrees.

3.1. **Using 2-transitive groups for special degrees.**

The results of this section are summarized in Examples 3.16. We begin with an integer $n \ge 3$, together with

- a group T acting transitively (though not necessarily faithfully) on the unordered pairs of distinct points in $\{1, \dots, n\}$ (i.e., T is 2-*homogeneous*),
- a presentation $\langle X \mid R\rangle$ of T,

- a subset X_1 of T such that $T_1 = \langle X_1 \rangle$ is the stabilizer of 1 (we usually have $X_1 \subset X$), and
- a word w in X that moves 1 and lies in A_n (when w is viewed inside T).

The last requirement implies that the permutation group $\bar{T}$ induced by T may not be inside A_n; when it is inside then the following lemma and its proof are somewhat simpler. Note that, if T is not 2-transitive, then $\bar{T}$ has odd order, and hence lies in A_n. (Here the order is odd because an involution in $\bar{T}$ would allow some *ordered* 2-set to be moved to any other one.)

Lemma 3.3. *If* $J = \langle X, z \mid R, z^3 = 1, (zz^w)^2 = 1, z^u = z^{\mathrm{sign}(u)}$ *for* $u \in X_1 \rangle$, *then* $J \cong A_{n+2} \times T$.

Proof. View T as a subgroup of $A_{n+2} = \mathrm{Alt}\,\{1, \dots, n, n+1, n+2\}$, with each $t \in T$ inducing $(n+1, n+2)^{\mathrm{sign}(t)}$ on $\{n+1, n+2\}$. Define $\varphi\colon X \cup \{z\} \to A_{n+2} \times T$ by

$$\begin{aligned} \varphi(x) &= (x(n+1, n+2)^{\mathrm{sign}(x)}, x) \quad \text{for } x \in X \\ \varphi(z) &= (z', 1) \quad \text{where} \quad z' = (1, n+1, n+2). \end{aligned} \tag{3.4}$$

Then the image of φ satisfies the defining relations for J, and we obtain a homomorphism $\varphi\colon J \to A_{n+2} \times T$. We claim that φ is a surjection. For, since T is 2-homogeneous on $\{1, \dots, n\}$, we have $\langle \varphi(z)^{\langle \varphi(X) \rangle} \rangle = A_{n+2} \times 1$. Here, A_{n+2} contains $\langle X \rangle = T$, while $\langle \varphi(X) \rangle$ projects onto T in the second component, so that $\varphi(J)$ also contains $1 \times T$. Hence, φ is, indeed, surjective.

Then there is also a surjection $\pi\colon J \to A_{n+2}$. By Lemma 2.1, J has a subgroup we identify with $T = \langle X \rangle$. We also view z as contained in J.

Since $\langle z \rangle^{T_1} = \langle z \rangle$ by our relations, we have $|\langle z \rangle^T| \le n$; but $|\pi(\langle z \rangle^T)| = n$ and so $|\langle z \rangle^T| = n$. Similarly, $|z^{T \cap A_n}| = n$. Consequently, T acts on $\langle z \rangle^T$ and $T \cap A_n$ acts on $z^{T \cap A_n}$ as they do on $\{1, \dots, n\}$.

Moreover, if T is not inside A_n, then our sign condition in the presentation implies that $|z^T| = 2n$, and $T \cap A_n$ has 2 orbits on z^T, namely, $z^{T \cap A_n}$ and $(z^{-1})^{T \cap A_n}$.

By 2-homogeneity, any unordered pair of distinct members of $\langle z \rangle^T$ is T-conjugate to $\{\langle z \rangle, \langle z \rangle^w\}$. If $T \cap A_n$ is 2-homogeneous, then any unordered pair of distinct members of z^T is $T \cap A_n$–conjugate to $\{z, z^w\}$. Since the relation $(zz^w)^2 = 1$ in the lemma implies that $(z^w z)^2 = 1$, it follows that z^T satisfies (3.1), so that $N := \langle z^T \rangle \cong A_{n+2}$.

If $T \cap A_n$ is not 2-homogeneous then, by hypothesis, T is 2-transitive but $T \cap A_n$ is not. We claim that we still have $N := \langle z^T \rangle \cong A_{n+2}$. For, any ordered pair of distinct members of $\langle z \rangle^T$ is $T \cap A_n$–conjugate to $(\langle z \rangle, \langle z \rangle^w)$ or to one other pair, $(\langle z \rangle, \langle z \rangle^y)$, say, where $y \in T \cap A_n$. Some $g \in T \backslash A_n$ satisfies $(\langle z \rangle, \langle z \rangle^w)^g = (\langle z \rangle, \langle z \rangle^y)$. Clearly, $z, z^w, z^y \in z^{T \cap A_n}$. Since $g \notin A_n$, it follows that both z^g and $(z^w)^g$ lie in the other $T \cap A_n$–class $(z^{-1})^{T \cap A_n}$ of z. Thus, $z^g = z^{-1}$ and $(z^w)^g = (z^y)^{-1}$, so that $(zz^y)^2 = ([z^{-1}(z^w)^{-1}]^2)^g = 1$ by our relations, and we again have $N \cong A_{n+2}$ by (3.1).

Clearly $N \trianglelefteq J$ and $J/N \cong T$. Then $|J| = |A_{n+2} \times T|$, so that $J \cong A_{n+2} \times T$. □

Examples 3.5. (1) Let p be an odd prime, $n = p+1$ and $T = \mathrm{SL}(2, p)$. Then T has a presentation with 2 generators and 2 relations [CR2], while T_1 can be generated by 2 elements. Thus, by the Lemma, $A_{p+3} \times \mathrm{SL}(2, p)$ has a presentation with 3 generators and 6 relations (cf. Examples 3.16(5) and 3.19(9)).

(2) We can do somewhat better by taking T to be $\mathrm{AGL}(1, p) = P \rtimes T_1$, with P cyclic of odd prime order p and T_1 cyclic of order $p - 1$. By [Neu], if r and s are

integers such that $\mathbb{F}_p^* = \langle r\rangle$ and $s(r-1) \equiv -1 \pmod p$, then

$$T = \mathrm{AGL}(1,p) = \langle a, b \mid a^p = b^{p-1}, (a^s)^b = a^{s-1}\rangle. \tag{3.6}$$

This time $|X_1| = 1$, and hence $A_{p+2} \times T$ *has a presentation with* 3 *generators and* 5 *relations.*

(3) An example of a 2-homogeneous group that is not 2–transitive is the subgroup $T = \mathrm{AGL}(1,p)^{(2)}$ of index 2 in $\mathrm{AGL}(1,p)$ for a prime $p \equiv 3 \pmod 4$, $p > 3$. This time $T = P \rtimes T_1$ with P cyclic of order p and T_1 cyclic of order $(p-1)/2$. By [Neu],

$$T = \mathrm{AGL}(1,p)^{(2)} = \langle a, b \mid a^p = b^{(p-1)/2}, (a^s)^b = a^{s-1}\rangle,$$

where this time $\mathbb{F}_p^{*2} = \langle r\rangle$ and $s(r-1) \equiv -1 \pmod p$. Once again $A_{p+2} \times T$ *has a presentation with* 3 *generators and* 5 *relations.*

(4) For future reference we note that, for any prime $p \equiv 3 \pmod 4$ with $p > 3$,

$$\mathrm{AGL}(1,p)^{(2)} \times \mathbb{Z}_2 = \langle a, b \mid a^{2p} = b^{(p-1)/2}, (a^s)^b = a^{s-2}\rangle,$$

where $s(r-1) \equiv -2 \pmod p$, and $\mathbb{F}_p^{*2} = \langle r\rangle$; since both s and $s+p$ both satisfy the preceding congruence, we may assume that with s odd. For, the presented group T surjects onto $\mathrm{AGL}(1,p)^{(2)} \times \mathbb{Z}_2$ by sending $a \to (a', z)$ and $b \to (b', 1)$, with a', b' playing the roles of a, b in (3), and $|z| = 2$. Since $a^{2ps} = (a^{2ps})^b = (a^{2p})^{s-2}$, we have $a^{4p} = 1$. Then $(a^{2p})^s = a^{2p}$, so that $(a^{ps})^b = a^{p(s-2)} = a^{-ps}$, and hence $a^{ps} = (a^{ps})^{b^{(p-1)/2}} = a^{-ps}$ since $(p-1)/2$ is odd. Now $a^{4p} = 1 = a^{2ps}$. Since s is also odd, it follows that $a^{2p} = 1$, and also that $\langle a\rangle \trianglelefteq T$. Then $a^p \in Z(T)$, and T is as claimed.

Corollary 3.7. *If p is prime then A_{p+2} has a presentation with* 3 *generators*, 6 *relations and bit-length $O(\log p)$.*

Proof. By Example 3.5(2), the lemma provides a presentation for $A_{p+2} \times \mathrm{AGL}(1,p)$ with 3 generators $\mathbf{a}, \mathbf{b}, \mathbf{z}$. Under the isomorphism (3.4) in the lemma, in $A_{p+2} \times T$ we have $\mathbf{b} = \big(b(p+1, p+2), b\big)$ for a $(p-1)$-cycle $b \in T_p$ (the point 1 in that lemma is now the point $p \equiv 0$ fixed by b). Then b^a moves 0 (and fixes $p+1$ and $p+2$), so that $\mathbf{b^a z} = \big(b^a(p+1, p+2), b^a\big)(z', 1) = (c, b^a)$ for a p-cycle c. Thus, $(\mathbf{b^a z})^p = (1, b^a)$. Since T is the normal closure of b^a in T, imposing the relation $(\mathbf{b^a z})^p = 1$ gives a presentation for A_{p+2}. □

Note that this is not, however, a short presentation (cf. Section 2).

For some choices of p we can improve the preceding result:

Corollary 3.8. *For any prime $p \equiv 11 \pmod{12}$,*

(i) *$A_{p+2} \times \mathrm{AGL}(1,p)^{(2)}$ has a presentation with* 2 *generators*, 3 *relations and bit-length $O(\log p)$; and*

(ii) *A_{p+2} has a presentation with* 2 *generators*, 4 *relations and bit-length $O(\log p)$.*

Proof. (i) Since $p \equiv 3 \pmod 4$, we can let $T = \mathrm{AGL}(1,p)^{(2)} \le A_p$, r and s be as in Example 3.5(3). Consider the group J defined by the presentation

$$\langle a, g \mid a^p = b^{(p-1)/2}, (a^s)^b = a^{s-1}, (zz^a)^2 = 1\rangle,$$

where $b := g^3$ and $z := g^{(p-1)/2}$.

First note that $A_{p+2} \times T$ satisfies this presentation. Namely, let T act on $\{1, \dots, p, p+1, p+2\}$, fixing $p+1$ and $p+2$. Let g be the product of the 3-cycle $(1, p+1, p+2)$ and an element of T having two cycles of length $(p-1)/2$

on $\{2,\dots,p\}$. Since $p \equiv 2 \pmod 3$, g has order $3(p-1)$, so that $T = \langle a, g^3\rangle$ satisfies the presentation in Example 3.5(3). The final relation $(zz^a)^2 = 1$ holds as in Lemma 3.3 using $w = a$.

By Example 3.5(3) and Lemma 2.1, J has a subgroup we can identify with $T = \langle a, b\rangle$. Clearly, $T_p = T_0 = \langle b\rangle$. The remaining relations $z^3 = 1$ and $z^b = z$ in Lemma 3.3 are automatic: they hold in $\langle g\rangle$. This proves (i).

(ii) This is similar to the preceding corollary. This time b has two cycles of length $(p-1)/2$, and we obtain

$$(3.9)\quad A_{p+2} \cong \langle a, g \,|\, a^p = g^{3\kappa}, (a^s)^{g^3} = a^{s-1}, \big(g^\kappa (g^\kappa)^a\big)^2 = \big(g^3 (g^\kappa)^a (g^\kappa)^{a^{-1}}\big)^{\kappa+1} = 1\rangle$$

with $\kappa = (p-1)/2$, $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^{*2} = \langle r\rangle$, since -1 is a non-square mod p. For, we view (i) as a presentation for $A_{p+2} \times \mathrm{AGL}(1,p)^{(2)}$ with generators $\mathbf{a}$ and $\mathbf{g}$, and we use $\mathbf{b} := \mathbf{g}^3$ and $\mathbf{z} := \mathbf{g}^\kappa$. By (3.4), as in the proof of Corollary 3.7 we have $\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}} = \big(b, b\big)(z'^a z'^{a^{-1}}, 1) = (c_1c_2, b)$ for disjoint cycles c_1 and c_2 of length $\kappa+1$, since $a(1)$ and $a(-1)$ are in different b-cycles. Then $(\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}})^{\kappa+1} = (1, b)$, and imposing the relation $(\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}})^{\kappa+1} = 1$ on the presentation in (i) gives (3.9). □

Remark 3.10. *In the presentations in the preceding corollaries, every cycle* $(k, k+1, \dots, l)$ (*with* $k-l$ *even*) *can be written as a word of bit-length* $O(\log p)$ *in the generators. Any even permutation with bounded support can also be expressed as word of bit-length* $O(\log p)$ *in the generators. For all of these elements, the indicated words use a bounded number of exponents.*

Namely, all 3-cycles $(k, p+1, p+2) = (1, p+1, p+2)^{a^{k-1}}$ have bit-length $O(\log p)$, therefore the same is true of any permutation of bounded support (because it is a product of a bounded number of such 3-cycles). In particular, if $x = (1,p)(p+1, p+2)$ then $xa = (1,\dots,p-1)(p+1,p+2)$ has bit-length $O(\log p)$. Since

$$(xa)^{-k}a^k = (p, 1, \dots, k)(p+1, p+2)^k,$$

if $k < p$ is even then the $(k+1)$-cycle $(p, 1, \dots, k)$ has bit-length $O(\log p)$, hence the same is true of all even cycles of the form $(k, \dots, l)$, $l < p$. The remaining cycles arise as, for example, $(k, \dots, p)(k, p, p+1) = (k+1, \dots, p, p+1)$.

Symmetric groups. There are analogous results for symmetric groups. This time we assume that our group $T = \langle X \mid R\rangle$ acts 2-transitively on $\{1, \dots, n\}$ and *does not lie in* A_n; let w be a word in X that moves 1 when w is viewed inside T. As above, the obvious examples are $\mathrm{AGL}(1,p)$ and $\mathrm{PGL}(2,p)$.

Lemma 3.11. *If* $J = \langle X, z \mid R, z^3 = 1, (zz^w)^2 = 1, [z, T_1] = 1\rangle$, *then* J *has a normal subgroup* $T \cap A_{n+2}$ *modulo which it is* S_{n+2}. *Moreover,* J *is isomorphic to a subgroup of index* 2 *in* $S_{p+2} \times T$ *that projects onto each factor.*

Proof. This time view T as a subgroup of S_{n+2} acting on $\{1, \dots, n, n+1, n+2\}$ but fixing $n+1$ and $n+2$. Then S_{n+2} acts on a set of size $n+2$, while T also acts on a set of size n. View $S_{n+2} \times T$ as acting on the disjoint union of these sets, and let H be the subgroup $A_{2n+2} \cap (S_{n+2} \times T)$ of index 2 (recall that T is not inside A_n).

This time we map J into H using a simpler version of (3.4): $\varphi(x) = (x, x)$ for $x \in X$. We identify $T = \langle X\rangle$ with a subgroup of J and z with an element of J. As before, T acts on z^T as it does on $\{1, \dots, n\}$, and hence is 2-transitive. Then the relation $(zz^w)^2 = 1$ and (3.1) imply that $N := \langle z^T\rangle \cong A_{n+2}$. Since $J/N \cong T$, we have $|J| = |A_{n+2}||T| = |H|$, so that $J \cong H$. □

Corollary 3.12. *Let p be a prime.*

(i) *S_{p+2} has a presentation with 3 generators, 6 relations and bit-length $O(\log p)$.*

(ii) *If $p \equiv 2 \pmod 3$ then the subgroup of index 2 in $S_{p+2} \times \mathrm{AGL}(1,p)$ that projects onto each factor has a presentation with 2 generators, 3 relations and bit-length $O(\log p)$.*

(iii) *If $p \equiv 2 \pmod 3$ then S_{p+2} has a presentation with 2 generators, 4 relations and bit-length $O(\log p)$.*

Proof. Part (i) follows from the preceding lemma together with Example 3.5(2), while (ii) is proved exactly as in Corollary 3.8 by using that example. This time in (iii) we obtain

$$
\begin{aligned}
S_{p+2} \cong \langle a, g \mid a^p = (g^3)^{p-1}, (a^s)^{g^3} = a^{s-1},& \\
\big(g^{p-1}(g^{p-1})^a\big)^2 = \big(g^6 (g^{p-1})^{a^{-1}} (g^{p-1})^{a^{-r}}\big)^{(p+1)/2} = 1\rangle&
\end{aligned}
\tag{3.13}
$$

with $s(r-1) \equiv -1$ (mod p and $\mathbb{F}_p^* = \langle r \rangle$. For, once again we view (ii) as a presentation with generators $\mathbf{a}$ and $\mathbf{g}$, we use $\mathbf{b} := \mathbf{g}^3$ and $\mathbf{z} := \mathbf{g}^{p-1}$, and then $\mathbf{b}^2 \mathbf{z}^{\mathbf{a}^{-1}} \mathbf{z}^{\mathbf{a}^{-r}} = \big(b^2, b^2\big)\big(z'^{a^{-1}} z'^{a^{-r}}, 1\big) = (c, b^2)$ for a $(p+1)/2$-cycle c. Factoring out the normal closure of $b^{p+1} = b^2$ in T, we obtain (3.13). □

Remark 3.14. *For the presentation in* Corollary 3.12, *the assertions in* Remark 3.10 *hold once again for even permutations. They also hold for odd permutations if $p \equiv 11 \pmod{12}$.*

Even permutations are handled as before. Odd permutations are slightly more delicate: they require constructing a transposition of the required bit-length, and we are only able to achieve this when $p \equiv 11 \pmod{12}$. First we recall a group-theoretic version of "Horner's Rule" [GKKL1, (3.3)] for elements v, f in any group:

$$
v v^f v^{f^2} \cdots v^{f^n} = (v f^{-1})^n v f^n
\tag{3.15}
$$

for any positive integer n.

Note that

$b_2 := \mathbf{b}^{(p-1)/2} = (1, p-1)(2, p-2) \cdots \big((p-1)/2, p-(p-1)/2\big)$

is an odd permutation since $p \equiv 3 \pmod 4$. We will use several additional permutations:

$c(i,j) := \mathbf{z}^{\mathbf{a}^{-i}} (\mathbf{z}^{\mathbf{a}^{-j}})^{-1} \mathbf{z}^{\mathbf{a}^{-i}} = (i,j)(p+1, p+2)$ for $1 \le i, j \le p$,

$v_\bullet := c(1, p-1) c(2, p-2) = (1, p-1)(2, p-2)$,

$c_{(p-1)/2} := \big(c(1,2)\mathbf{a}\big)^{(p-1)/2-2} c(1,2) \mathbf{a}^{-((p-1)/2-2)} = \big(1, 2, \ldots, (p-1)/2\big)$

since $(p-1)/2$ is odd (cf. (3.15)),

$c_\bullet := c_{(p-1)/2}\, c_{(p-1)/2}^{-\mathbf{a}^{(p+1)/2}} = \big(1, 2, \ldots, (p-1)/2\big)\big(p-1, p-2, \ldots, p-(p-1)/2\big)$,

and

$$
\begin{aligned}
v := \big(c(1,p-1)c_\bullet^{-1}\big)^{(p-1)/2} c(1,p-1) c_\bullet^{(p-1)/2} &= \big(c(1,p-1)c_\bullet^{-1}\big)^{(p-1)/2} c(1,p-1) \\
&\equiv (1,p-1)(2,p-2)\cdots\big((p-1)/2, p-(p-1)/2\big)(p+1,p+2) \quad \text{(cf. (3.15))}.
\end{aligned}
$$

Then $v b_2$ is the transposition $(p+1, p+2)$, as required (compare Section 3.5).

We do not know if the final assertion in the preceding remark holds when $p \equiv 1 \pmod 4$.

Examples 3.16. *We summarize versions of the previous presentations for special values of n.* Additional explicit presentations appear in Examples 3.19.

Consider a prime $p > 3$.

(1) $A_{p+2} = \langle a, b, z \mid a^p = b^{p-1}, (a^s)^b = a^{s-1}, z^3 = (zz^a)^2 = 1, z^b = z^{-1}, (b^a z)^p = 1\rangle$, where $s(r-1) \equiv -1 \pmod p$ with $\mathbb{F}_p^* = \langle r\rangle$ (by Corollary 3.7).

(2) A_{p+2} for $p \equiv 11 \pmod{12}$: see (3.9).

(3) S_{p+2} for $p \equiv 2 \pmod 3$: see (3.13).

(4) $S_{p+2} = \langle a, b, z \mid a^p = b^{p-1}, (a^s)^b = a^{s-1}, z^3 = (zz^a)^2 = [z, b] = 1, (b^2 z^{a^{-1}} z^{a^{-r}})^p = 1\rangle$ for $p \equiv 1 \pmod 3$, where $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^* = \langle r\rangle$ (using Corollary 3.12).

(5) We will give several presentations of A_{p+3} both here and in Example 3.19(9). Let $\mathbb{F}_p^* = \langle j\rangle$ and $j\bar{j} \equiv 1 \pmod p$. Then

$$\begin{aligned} A_{p+3} = \langle x, y, z \mid x^2 = (xy)^3, (xy^4xy^{(p+1)/2})^2 y^p x^{2[p/3]} = 1, \\ z^3 = (zz^x)^2 = [y, z] = [h, z] = 1, (hz^{(xy)^{-1}} z^{(xy^j)^{-1}})^{(p+1)/2} = 1\rangle, \end{aligned}$$

where we have abbreviated $h := y^{\bar{j}}(y^j)^x y^{\bar{j}} x^{-1}$. This uses the following presentation for $T := \mathrm{SL}(2, p)$, obtained in [CR2] using [Sun]:

$$\mathrm{SL}(2, p) = \langle x, y \mid x^2 = (xy)^3, (xy^4xy^{(p+1)/2})^2 y^p x^{2[p/3]} = 1\rangle, \tag{3.17}$$

where x and y arise from elements of order 4 and p, respectively. (These correspond to the matrices t and u given later in (4.4).) Then $T_1 = \langle X_1\rangle$ with $X_1 := \{y, h\}$ in the notation used in Lemma 3.3; the final relation in the presentation for A_{p+3} is obtained as in the proof of Corollary 3.8(ii).

$$\begin{aligned} A_{p+3} = \langle u, h, t, z \mid u^p = t^2 = 1, u^h = u^{j^2}, h^t = h^{-1}, t = uu^tu, ht = u^{\bar{j}}(u^j)^t u^{\bar{j}}, \\ z^3 = (zz^t)^2 = [u, z] = [h, z] = 1, (hz^{(tu)^{-1}} z^{(tu^j)^{-1}})^{(p+1)/2} = 1\rangle. \end{aligned}$$

This uses Lemma 3.3 together with the presentation for $T := \mathrm{PSL}(2, p)$ given in [GKKL1, Theorem 3.6]. A similar presentation can be obtained using the presentation for $\mathrm{PSL}(2, p)$ in [To].

(6) Once again let $\mathbb{F}_p^* = \langle j\rangle$. Then

$$\begin{aligned} S_{p+3} = \langle u, h, t, z \mid u^p = t^2 = 1, u^h = u^j, h^t = h^{-1}, t = uu^tu, \\ z^3 = (zz^t)^2 = [u, z] = [h, z] = 1, (hz^{tu})^{p+1} = 1\rangle. \end{aligned}$$

This uses Lemma 3.3 together with a presentation $\langle u, h, t \mid u^p = t^2 = 1, u^h = u^j, h^t = h^{-1}, t = uu^tu\rangle$ for $T := \mathrm{PGL}(2, p)$ analogous to [GKKL1, Theorem 3.6]. The final relation is obtained as in the proof of Corollary 3.8(iii). Once again one could also use [To] for a presentation of T .

3.2. **Small** n**.** In order to handle a few degrees $n < 50$ we will need further variations on the idea used in Corollaries 3.8 and 3.12. All of the general presentations below have bit-length $O(\log n)$, but this is not significant since our goal involves bounded n. We suspect that most readers will wish to skip this section.

In Table 1 we summarize the cases needed later. For this table and our variation on Corollaries 3.8 and 3.12 we use the following notation:

- T is a group acting transitively (though not necessarily faithfully) on $\{1, \ldots, n\}$.
- T has exactly ρ orbits of unordered pairs of distinct points.
- $T = \langle X \mid R\rangle$.
- $T_1 = \langle X_1\rangle$, where T_1 is again the stabilizer of 1.
- $x_1 \in X_1 \cap X$ has order k not divisible by 3.

Table 1. Some small n

G	n	T	$\lvert R\rvert$	ρ	$\lvert X_1\rvert$	$\lvert x_1\rvert$	gens	rels	in Ex.♯
S_{11}	$9+2$	$\mathrm{AGL}(1,9)$	4	1	1	8	2	6	3.19(1)
A_{11}	$9+2$	$\mathrm{PSL}(2,8)$	2	2	1	4	2	5	3.19(2)
A_{11}	$9+2$	$\mathrm{AGL}(1,9)^{(2)}$	4	2	1	4	2	7	3.19(3)
S_{12}	$10+2$	$\mathrm{PGL}(2,9)$	3	1	2	8	2	6	3.19(4)
A_{12}	$10+2$	$6.\mathrm{PSL}(2,9)$	2	1	2	8	2	5	3.19(5)
A_{23}	$21+2$	12.PSL(3,4)	2	1	2	5	2	5	3.19(6)
A_{23}	$\binom{7}{2}+2$	$6.A_7$	2	2	2	5	2	6	3.19(7)
A_{24}	$22+2$	$12.M_{22}$	2	1	2	7	2	5	3.19(8)
A_{24}	$2\cdot 11+2$	$\mathrm{AGL}(1,11)^{(2)}\times\mathbb{Z}_2$	2	3	1	5	2	6	3.19(12)
A_{24}	$2\cdot 11+2$	$\mathrm{AGL}(1,11)$	2	4	1	5	2	7	3.19(13)
A_{47}	$\binom{10}{2}+2$	A_{10}	2	2	2	8	2	6	3.19(10)
A_{47}	$\binom{10}{2}+2$	$A_{10}\times\mathrm{SL}(2,7)$	4	2	2	7	2	8	3.19(11)
A_{48}	$2\cdot 23+2$	$\mathrm{AGL}(1,23)^{(2)}\times\mathbb{Z}_2$	2	3	1	11	2	6	3.19(12)
A_{48}	$2\cdot 23+2$	$\mathrm{AGL}(1,23)$	2	4	1	11	2	7	3.19(13)

- T is also viewed as a subgroup of $\mathrm{Alt}\,\{1,\dots,n+2\}$.

Thus, $\langle T^{S_{n+2}}\rangle$ is A_{n+2} if T is in A_{n+2}, and S_{n+2} otherwise.

Proposition 3.18. *If T is the normal closure of one of its elements, then $\langle T^{S_{n+2}}\rangle$ has a presentation with $|X|$ generators and $|R|+\rho+|X_1|$ relations.*

If $T\le A_n$ then $A_{n+2}\times T$ has a presentation with $|X|$ generators and $|R|+\rho+|X_1|-1$ relations.

Proof sketch. Let $w_1,\dots,w_\rho$ be words in X such that the ρ pairs $\{1,w_i^{-1}(1)\}$ are in different T-orbits. Let $X=\{x_1,\dots\}$ and $R=\{r_1,\dots\}$ with each r_i a word $r_i(x_1,\dots)$ in X. Let $X':=X\backslash\{x_1\}$ and $X_1':=X_1\backslash\{x_1\}$. We claim that

$$\begin{aligned} J:=\langle X',g \mid r_i(g^3,\dots)=1 \text{ for all } i,&\\ [g^k,X_1']=\big(g^k(g^k)^{w_j}\big)^2=1 \text{ for all } j\rangle& \end{aligned}$$

is isomorphic to $H:=A_{2n+2}\cap(S_{n+2}\times T)$. For, view $S_{n+2}\times T$ as acting on the disjoint union $\{1,\dots,n,n+1,n+2\}\,\dot\cup\,\{1,\dots,n\}$ with T fixing $n+1$ and $n+2$. Let $z':=(1,n+1,n+2)$ and let the integer ν satisfy $3\nu\equiv 1 \pmod k$, so that $g:=x_1^\nu z'$ satisfies $g^3=x_1$. Then J surjects onto H as in the proof of Lemma 3.3.

We can identify $\tilde T:=\langle g^3,X'\rangle$ with a subgroup of J that is a homomorphic image of our original T. Our hypotheses guarantee that $z:=g^k$ commutes with X_1. Then $|z^{\tilde T}|=n$, and we can use (3.1) as before. □

Examples 3.19(12) and 3.19(13) contain further variations on the idea behind the Proposition.

Examples 3.19. (1) $n=11$: $T=\mathrm{AGL}(1,9)$ has the presentation $\langle a,b\mid a^3=b^8=1, a^{b^2}=aa^{-b},[a,a^b]=1\rangle$, $\rho=|X_1|=1$ and $|x_1|=8$, so that S_{11} *has a presentation with* 2 *generators and* $4+1+1$ *relations.* However, for use in Theorem C it is easier simply to use the presentation of S_{11} with 2 generators and 6 relations in [Ar, p. 54] (cf. [CoMo, p. 64]). See Remark 4.8 for a generalization.

(2) $n = 11$: $T = \mathrm{PSL}(2,8)$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho = 1, |X_1| = 2$ and $|x_1| = 7$, so that Proposition 3.18 produces a *presentation of* A_{11} *with* 2 *generators and* $2+2+1$ *relations.* Once again, it has long been known that A_{11} has a presentation with 2 generators and 6 relations [CoMo, p. 67].

(3) $n = 11$: T has index 2 in $\mathrm{AGL}(1,9)$, T has the presentation $\langle a, b \mid a^3 = b^4 = 1, a^{b^2} = a^{-1}, [a, a^b] = 1\rangle$, $\rho = 2$, $|X_1| = 1$ and $|x_1| = 4$, so that A_{11} *has a presentation with* 2 *generators and* $4+2+1$ *relations.* See Remark 4.8 for a generalization.

(4) $n = 12$: $T = \mathrm{PGL}(2,9)$ has presentations with 2 generators and 3 relations provided by G. Havas [Hav]. The following are some of his many presentations related to A_6:

$$\begin{aligned}
\mathrm{PGL}(2,9) &= \langle a, b \mid b^5 = (aba)^2 = ab^{-1}a^4ab^{-1}a^3b^{-1}ab^2 = 1\rangle\\
\mathrm{PGL}(2,9) &= \langle a, b \mid b^8 = baba^3bab^2 = b^{-1}a^{-1}b^{-2}a^2ba^2b^{-1} = 1\rangle\\
S_6 &= \langle a, b \mid a^{-1}b^{-1}a^3b^{-1}a^{-2} = (ab)^5 = b^{-3}a^{-1}b^{-1}a^2b^{-2}aba = 1\rangle\\
S_6 &= \langle a, b \mid a^4 = b^{-1}ab^{-2}a^2b^2a^{-1}b^{-1} = aba^{-1}b^{-1}a^{-2}b^{-1}aba^{-1}b = 1\rangle.
\end{aligned}$$

This time $\rho = 1$, $|X_1| = 2$ and we may assume that $|x_1| = 8$, so that S_{12} *has a presentation with* 2 *generators and* $3+1+2$ *relations.* Once again, it has long been known that S_{12} has a presentation with 2 generators and 7 relations in [Ar, p. 54] (cf. [CoMo, p. 64]).

(5) $n = 12$, $T \cong 6.\mathrm{PSL}(2,9) \cong 6.A_6$ has a presentation with 2 generators and 2 relations [Ro], $\rho = 1$, $|X_1| = 2$ and $|x_1| = 8$, so that A_{12} *has a presentation with* 2 *generators and* $2+1+2$ *relations.* Once again, it has long been known that A_{12} has a presentation with 2 generators and 7 relations [CoMo, p. 67].

(6) $n = 23$: $T = 12.\mathrm{PSL}(3,4)$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho = 1, |X_1| = 2$ and $|x_1| = 5$, so that Proposition 3.18 produces a *presentation of* A_{23} *with* 2 *generators and* $2+2+1$ *relations.*

(7) $n=23$: $T = 6.A_7$ has a presentation with 2 generators and 2 relations [CRKMW], $n = \binom{7}{2}$, $\rho = 2$, $|X_1| = 2$ and $|x_1| = 5$, so that A_{23} *has a presentation with* 2 *generators and* $2+2+2$ *relations.*

(8) $n = 24$: $T = 12.M_{22}$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho = 1, |X_1| = 2$ and $|x_1| = 7$, so that Proposition 3.18 produces a presentation of A_{24} *with* 2 *generators and* $2+2+1$ *relations.*

(9) *If* $p > 3$ *is prime then* $A_{p+3} \times \mathrm{SL}(2,p)$ *has a presentation with* 2 *generators and* 4 *relations.* Namely, apply Proposition 3.18 using (3.17) and $|R| = 2 = |X_1|$, $\rho = 1$, $x_1 = y$. It follows that A_{p+3} *has a presentation with* 2 *generators and* 5 *relations.* We will need this below in (11).

Explicitly, as in Example 3.16(5) we have

$$\begin{aligned}
A_{p+3} = \langle x, g \mid x^2 = (xg^3)^3, (xg^{12}xg^{3(p+1)/2})^2g^{3p}x^{2[p/3]} = 1,&\\
[g^p, h] = \big(g^p(g^p)^x\big)^2 = 1, (h(g^p)^{xg^3})^{p+1} = 1\rangle,&
\end{aligned}$$

where $h := g^{3\bar{\jmath}}(g^{3j})^x g^{3\bar{\jmath}}x^{-1}$ with $\mathbb{F}_p^* = \langle j\rangle$ and $j\bar{\jmath} \equiv 1 \pmod{p}$.

(10) $n = 47$: $T = A_{10}$ has the following presentation with 2 generators and 3 relations [Hav]:

$$A_{10} = \langle a, b \,|\, a^3b^{-1}ab^{-1}a^3ba^2b = a^2b^{-1}a^5b^{-3}a^3 = a^{-2}bab^{-1}aba^3bab^{-1}aba^{-2}b^{-1} = 1\rangle,$$

with $|a| = 15, |b| = 12$ and $|ab| = 8$; hence we modify this presentation so that the generators are a and ab. View T as acting on $\binom{10}{2}$ unordered pairs with $\rho = 2$, $|X_1| = 2$ and $x_1 = ab$, so that Proposition 3.18 produces a *presentation of* A_{45+2} *with* 2 *generators and* $2+2+2$ *relations.*

(11) *If* $p > 3$ *is prime then* $A_{\binom{p+3}{2}+2}$ *has a presentation with* 2 *generators and* 8 *relations.* For, let $T = A_{p+3} \times \mathrm{SL}(2, p)$ act on $\binom{p+3}{2}$ unordered pairs of a set of size $p+3$, with $\mathrm{SL}(2, p)$ acting trivially. Apply Proposition 3.18 using (9), with $|X| = 3$, $|R| = 4$, $\rho = 2$, $|X_1| = 2$ and $x_1 = y$. (Note that x_1 fixes $p + 2$ and $p + 3$ in (9).)

There is a similar presentation for $A_{\binom{p+2}{2}+2}$.

(12) *For any prime* $p \equiv 11 \pmod{12}$, A_{2p+2} *has a presentation with* 2 *generators and* 6 *relations.* We will vary the argument in Proposition 3.18 (and Lemma 3.3), using the transitive subgroup $T := \mathrm{AGL}(1, p)^{(2)} \times \langle t\rangle$ of the transitive group $\mathrm{AGL}(1, p) \times \langle t\rangle$ of degree $2p$, where t is an involution interchanging two blocks of size p. Note that the stabilizer of a point is cyclic of odd order $(p - 1)/2$. Moreover, T has $\rho = 3$ orbits of unordered pairs of the $2p$-set, with orbit-representatives as follows: contained in a block, or of the form $\{p, t(p)\}$, or of the form $\{p, t(1)\}$ (since t interchanges $\{p, t(1)\}$ and $\{1, t(p)\}$).

We view T as a subgroup of A_{2p+2} preserving the two new points $2p + 1$ and $2p + 2$, with t interchanging these points.

Let

$$J := \langle a, g \mid a^{2p} = b^{(p-1)/2}, (a^s)^b = a^{s-2}, (zz^{\mathrm{sign}(w_i)w_i})^2 = 1 \ (i = 1, 2, 3)\rangle$$

where $s(r - 1) \equiv -2 \pmod{p}$ with s odd and r of order $(p - 1)/2 \pmod{p}$, $b := g^3$, $z := g^{(p-1)/2}$, and suitable words $w_1, w_2, w_3 \in T$ (such that the pairs $\{z, z^{w_i}\}$ are in different T-orbits on the $2p$-set z^T); here sign refers to the behavior on the $2p$ points. For example, $\{z, z^t\}$, $\left\{z, z^{a^2}\right\}$ and $\left\{z, z^{a^2t}\right\}$ are representatives of these T-orbits. As in the proof of Proposition 3.18, using Example 3.5(4) we see that $A_{2p+2} \times T$ satisfies our presentation.

As usual, using Example 3.5(4) we can view T as the subgroup $\langle a, b\rangle$ of J. Exactly as in Lemma 3.3 (and Proposition 3.18), $|z| = 3$, $|z^T| = 2p$, and $(zz^g)^2$ for all $g \in \mathrm{AGL}(1, p)^{(2)}$ with $z^g \neq z$ while $(zz^{-gt})^2 = 1$ for all $g \in \mathrm{AGL}(1, p)^{(2)}$.

Then $N := \langle z^T\rangle \cong A_{2p+2}$ by (3.1), and hence $J = NT \cong A_{2p+2} \times \mathrm{AGL}(1, p)^{(2)} \times \langle t\rangle$. One further relation gives a presentation of A_{2p+2} with 3 generators and $2 + 3 + 1$ relations.

Explicitly:

$$\begin{aligned} A_{2p+2} = \langle a, g \mid a^{2p} = b^{(p-1)/2}, (a^s)^b = a^{s-2},& \\ (zz^{-t})^2 = (zz^{a^2})^2 = (zz^{-a^2t})^2 = 1, (tbz^{a^{-1}}z^a z^{ta^{-1}}z^{ta})^p = 1\rangle,& \end{aligned}$$

with z, r and s as above and once again $b := g^3$, $z := g^{(p-1)/2}$ and $t := a^p$, where the last relation is obtained as in the proof of Corollary 3.8(ii).

(13) *For any odd prime* $p \equiv 2 \pmod 3$, A_{2p+2} *has a presentation with* 2 *generators and* 7 *relations.* One difference between this example and the preceding one is that we now handle the case $p \equiv 1 \pmod 4$ using $T = \mathrm{AGL}(1,p)$.

First note that T acts transitively on a set of size $2p$, with cyclic point stabilizer of order $(p-1)/2$ and $\rho = 4$ orbits on unordered pairs of points. We again view T as a subgroup of A_{2p+2} preserving the additional points $2p+1$ and $2p+2$. Once again signs will refer to the actions of elements of T on the $2p$-set.

We replace the presentation of T in Example 3.5(1) by

$$T = \langle x, b \mid (xb^{-2})^p = b^{p-1},\ ((xb^{-2})^s)^b = (xb^{-2})^{s-2} \rangle,$$

with $x := ab^2 \in T$ of order $(p-1)/2$ fixing the point $r' := (1-r^2)^{-1}$, and r and s as in Example 3.5(1).

Consider the group

$$J := \langle g, b \,|\, (xb^{-2})^p = b^{p-1}, ((xb^{-2})^s)^b = (xb^{-2})^{s-2}, (zz^{\mathrm{sign}(w_i)w_i})^2 = 1\, (i = 1,2,3,4) \rangle,$$

where $x := g^3$, $z := g^{(p-1)/2}$, and $\{r', w_i^{-1}(r')\}$, $1 \le i \le 4$, are representatives for the orbits of T on pairs of the $2p$ points. Then J surjects onto $A_{2p+2} \times T$, and we can view $T = \langle x, b \rangle \le J$. In particular, $|g| = 3(p-1)/2$ and so $z^3 = 1$.

Since x centralizes z, as usual we obtain $|z^T| = 2p$ and $N := \langle z^T \rangle \cong A_{2p+2}$ by (3.1), and then $J \cong A_{2p+2} \times T$. One further relation produces the desired presentation.

3.3. **Gluing alternating and symmetric groups.** We now turn to the case of all alternating and symmetric groups, starting with a general gluing lemma:

Lemma 3.20. *Let* $G = \langle X \mid R \rangle$ *and* $\bar{G} = \langle \bar{X} \mid \bar{R} \rangle$ *be presentations of* S_n *and* S_m*, respectively, and let* $m, n > k \ge l+2 \ge 4$*. Consider embeddings* $\pi\colon G \to S_{m+n-k}$ *and* $\bar{\pi}\colon \bar{G} \to S_{m+n-k}$ *into* $\mathrm{Sym}\,\{-m+k+1, \dots, n\}$ *such that*

$$\pi(G) = \mathrm{Sym}(\{1, \dots, n\}) \quad \textit{and} \quad \bar{\pi}(\bar{G}) = \mathrm{Sym}(\{-m+1+k, \dots, k\}).$$

Suppose that $a, b, c, d \in G$ *and* $\bar{a}, \bar{b}, \bar{c}, \bar{e} \in \bar{G}$*, viewed as words in* X *or* $\bar{X}$*, respectively, are nontrivial permutations such that*

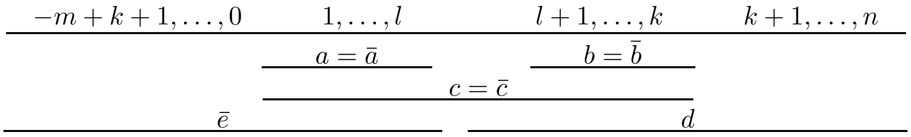
- $\pi(a) = \bar{\pi}(\bar{a}) \in \mathrm{Sym}(\{1, \dots, l\}) < \pi(G) \cap \bar{\pi}(\bar{G})$,
- $\pi(b) = \bar{\pi}(\bar{b}) \in \mathrm{Sym}(l+1, \dots, k) < \pi(G) \cap \bar{\pi}(\bar{G})$,
- $\pi(c) = \bar{\pi}(\bar{c}) \in \mathrm{Sym}(\{1, \dots, k\}) < \pi(G) \cap \bar{\pi}(\bar{G})$,
- $\pi(d) \in \mathrm{Sym}(\{l+1, \dots, n\}) < \pi(G)$,
- $\bar{\pi}(\bar{e}) \in \mathrm{Sym}(\{-m+1+k, \dots, l\}) < \bar{\pi}(\bar{G})$,
- $\langle \pi(a), \pi(c) \rangle = \mathrm{Sym}(\{1, \dots, k\}) < \pi(G) \cap \bar{\pi}(\bar{G})$,
- $\langle \pi(b), \pi(d) \rangle = \mathrm{Sym}(\{l+1, \dots, n\}) < \pi(G)$, *and*
- $\langle \bar{\pi}(\bar{a}), \bar{\pi}(\bar{e}) \rangle = \mathrm{Sym}(\{-m+1+k, \dots, l\}) < \bar{\pi}(\bar{G})$.

Then

$$J = \langle X, \bar{X} \mid R, \bar{R}, a = \bar{a}, c = \bar{c}, [d, \bar{e}] = 1 \rangle \tag{3.21}$$

is a presentation of $S_{m+n-k} = \mathrm{Sym}\,\{-m+k+1, \dots, n\}$*, where* $\langle X \rangle = S_n$ *acts on* $\{1, \dots, n\}$ *and* $\langle \bar{X} \rangle = S_m$ *acts on* $\{-m+1+k, \dots, k\}$*.*

The following picture might be helpful.



Proof. The restrictions on m, n, k and l are designed to guarantee that the desired permutations exist. There is a surjection $J \to S_{m+n-k}$ (note that our 3 extra relations are satisfied). By Lemma 2.1, J has subgroups we identify with $G = \langle X\rangle = S_n$ and $\bar G = \langle \bar X\rangle = S_m$.

By our relations, $a = \bar a$ and $c = \bar c$. Then the assumption $\pi(b) = \bar\pi(\bar b)$ states that b and $\bar b$ represent the same element of $\langle a, c\rangle = \langle \bar a, \bar c\rangle$, so that the additional relation $b = \bar b$ is forced to hold in J. Then we also have the following relations:

- $[d, \bar a] = [d, a] = 1$ because $d, a \in G = S_n$ have disjoint supports,
- $[d, \bar e] = 1$ by the last relation in the presentation (3.21),
- $[b, \bar a] = [b, a] = 1$ because $b, a \in G$ have disjoint supports, and
- $[b, \bar e] = [\bar b, \bar e] = 1$ because $\bar b, \bar e \in \bar G$ have disjoint supports.

Therefore

$$[\langle b, d\rangle, \langle \bar a, \bar e\rangle] = 1, \tag{3.22}$$

where $\langle b, d\rangle = \mathrm{Sym}(\{l+1,\dots,n\})$ and $\langle \bar a, \bar e\rangle = \mathrm{Sym}(\{-m+1+k,\dots,l\})$.

The symmetric groups G and $\bar G$ are generated, respectively, by the $n-1$ and $m-1$ transpositions $x_i := (i, i+1)$, $1 \le i < n$, and $x_i := \overline{(i, i+1)}$, $-m+1+k \le i < k$. The identification of the two copies of $S_k = \langle a, c\rangle = \langle \bar a, \bar c\rangle$ in (3.21) identifies the transpositions x_i, $1 \le i < k$, common to these generating sets, producing a generating set of J consisting of $m+n-k-1$ involutions. These involutions satisfy the relations in the Coxeter presentation [Moo]

$$S_{m+n-k} = \langle x_i, -m+1+k \le i < n \mid x_i^2 = (x_ix_{i+1})^3 = (x_ix_j)^2 = 1 \\ \text{for all possible } i, j \text{ with } j-i \ge 2\rangle:$$

any two x_i either both lie in G, or both lie in $\bar G$, or they commute by (3.22) since one is in $\langle b, d\rangle$ and the other is in $\langle \bar a, \bar e\rangle$. □

There is a great deal of flexibility in the choice of the elements $a, b, c, d, \bar a, \bar b, \bar c, \bar e$. In the proof of Theorem 3.36 we will need to require that $|a| = 3$, which is possible provided that $l \ge 3$.

The last three relations in (3.21) are similar to ones used in the proof of [GKKL1, Theorem 3.17]; and they are used both there and here in essentially the same manner. However, the situation in that paper was more delicate, due to length considerations: the preceding presentation does not even have bit-length $O(\log n)$. We will deal with this requirement when we use this lemma in the proof of Theorems 3.27 and 3.36.

Lemma 3.23. *A presentation of A_{m+n-k} is obtained as in the preceding lemma by replacing symmetric groups by alternating groups throughout* (3.21) *and assuming that $m, n > k \ge l+3 \ge 6$.*

Proof. Once again, the restrictions on m, n, k and l are designed to guarantee that the desired permutations exist. The previous picture can again be used. As in the proof of the preceding lemma, (3.21) implies that (3.22) holds.

We will use the presentation (3.1) with the union of the the generating sets

$x_i := (1,2,i)$ for G and $x_j := \overline{(1,2,j)}$ for $\bar{G}$, where $3 \le i \le n$ and $-m+1+k \le j \le k$ but $j \ne 1,2$. As above, (3.21) implies that $(1,2,i) = \overline{(1,2,i)}$ if $3 \le i \le k$.

Then all required relations in (3.1) are obvious except for

$$\big((1,2,i)\overline{(1,2,j)}\big)^2 = 1 \text{ with } k < i \le n \text{ and } -m+1+k \le j \le 0.$$

Recall that $\langle b,d\rangle = \mathrm{Alt}(\{l+1,\dots,n\}) < G = \mathrm{Alt}(\{1,\dots,n\})$, where $6 \le l+3 \le k < i \le n$. Then some $g \in \langle b,d\rangle$ sends k to i and fixes 1 and 2. By (3.22), g commutes with $\overline{(1,2,j)} \in \langle \bar{a},\bar{e}\rangle$. Consequently,

$$\big[\big((1,2,i)\overline{(1,2,j)}\big)^2\big]^g = \big((1,2,k)\overline{(1,2,j)}\big)^2 = \big(\overline{(1,2,k)}\,\overline{(1,2,j)}\big)^2 = 1,$$

as required in (3.1). □

Corollary 3.24. *If A_m has a presentation with M relations and if $m > k \ge 6$, then A_{2m-k} has a presentation with $M+4$ relations. The same holds for the corresponding symmetric groups using the weaker assumption $m > k \ge 4$.*

Proof. Let $G = \langle X \mid R\rangle$ be a presentation for A_m with M relations. In Lemma 3.23 we use $m = n$ and $l = 3$, but this time we introduce an additional generator y corresponding to an even permutation sending $\{1,\dots,m\} \to \{-m+1+k,\dots,k\}$ and inducing the identity on $\{1,\dots,k\}$.

Consider the group

$$J := \langle X, y \mid R, a = a^y, c = c^y, [d, e^y] = 1\rangle, \tag{3.25}$$

with $a,b,c,d,\bar{a} := a^y, \bar{b} := b^y, \bar{c} := c^y, \bar{e} := e^y$ playing the same roles as in Lemma 3.23. By that lemma with $\bar{X} := X^y$ and $\bar{R} := R^y$, J has a subgroup $K := \langle X, X^y\rangle \cong A_{2m-k}$.

Finally, we add an extra relation to ensure that our generator y is in K and that, as an element of A_{2m-k}, the action of y on $\{-m+k+1,\dots,m\}$ is as described above.

The group S_{2m-k} is dealt with in the same manner. □

Remark 3.26. We have just glued two subgroups A_m in order to obtain a group A_{2m-k}, or two subgroups S_m in order to obtain a group S_{2m-k}, in each case with suitable restrictions on m and k. There is a variation on this process that glues two subgroups S_m in order to obtain a group A_{2m-k} (view S_m as lying in A_{m+2}, as was done on occasion in Section 3.1).

3.4. **All alternating and symmetric groups.**

Before proving Theorem C, we begin with a weaker result:

Proposition 3.27. *For all $n \ge 5$, A_n and S_n have presentations with 3 generators and 10 relations.*

Proof. If $n \le 10$ then $n = p+2$ or $p+3$ for a prime p, and we have already obtained a presentation with fewer relations than required. By Ramanujan's version of Bertrand's Postulate [Ra], in all other cases we can write $n = 2p+4-k$ for a prime p and an integer k such that $m := p+2 > k \ge 6$, and then use Corollaries 3.9 and 3.23. (A related use of Bertrand's Postulate appears in [GKKL1, Theorem 3.9].) □

This proposition is weaker than Theorem C in two significant ways: the number of relations is larger than in that theorem, and bit-length is not mentioned. We deal with the second of these as follows:

Lemma 3.28. *In* Corollary 3.24 *and* Proposition 3.27, *it is possible to choose* a, c, d, e *such that* a *is a* 3*-cycle and the resulting presentation has bit-length* $O(\log n)$ *with a bounded number of exponents, each of which is at most* n, *if either*

(i) *the group is* A_n, *or*

(ii) *the group is* S_n *and we used a prime* $p \equiv 11 \pmod{12}$.

Proof. In Lemma 3.20 and Corollary 3.24 we can choose each of the elements a, c, d, e to be a product of a cycle of the form $(i, \dots, j)$ with $i < j$ and a permutation having bounded support (in fact, in Section 3.5 each will be chosen to be a cycle). We require a to be a 3-cycle (this is needed in Theorem 3.36); and then in the symmetric group case we will need c and e to be odd permutations (cf. the hypotheses of Lemma 3.20).

By Remark 3.10, when we glue two copies of A_{p+2} using relations of bit-length $O(\log n)$, the bit-length and exponents are as required, except perhaps for the crucial additional relation expressing y as word in $X \cup X^y$.

We now consider y; there is a reasonable amount of flexibility in the choice of y in Corollary 3.24 (recall that $m = p + 2$). In that corollary we were permuting the $n = 2p + 4 - k$ points

$$-p-1+k, -p+k, \cdots -1, 0;\ 1, \dots, k;\ k+1, k+2, \dots, p+1, p+2, \tag{3.29}$$

where we have alternating or symmetric groups on the first and last $p + 2$ points, with an A_k or S_k on the overlap.

If $p - k$ is even then we choose y to be the following product of $p + 2 - k$ transpositions:

$$y := (-p-1+k, p+2)(-p+k, p+1) \cdots (-1, k+2)(0, k+1). \tag{3.30}$$

We use the following additional permutations:

- $x := (-1, k+2)(0, k+1) = [(1, k+2)(2, k+1)]^{(-1,1)(0,2)}$, which we have written using permutations from the two alternating groups, and
- $u^{-1} := (1, \dots, k, k+1, \dots, p+2)(1, \dots, k, 0, -1, \dots, -p+k-1)$
  $= (1, \dots, k, k+1, \dots, p+2)(1, \dots, k, k+1, \dots, p+2)^y$.

By Remarks 3.10 and 3.14, u and hence also s can be expressed as a word of bit-length $O(\log p)$ in $X \cup X^y$ using a bounded number of exponents; then so can

$$y = xx^{u^2}x^{u^4} \cdots x^{u^{p-k}} = (xu^{-2})^{(p-k)/2}xu^{p-k}, \tag{3.31}$$

using (3.15).

If $p - k$ is odd let $v := (-p-1+k, p+2)(-p+k, p+1) \cdots (-3, k+4)$ and use

$$\begin{aligned} y :=\ & v(-2, k+3)(-1, k+2, 0, k+1) \\ =\ & v[(2, k+3)(1, k+2, k, k+1)]^{(2,-2)(1,-1)(0,k)}. \end{aligned} \tag{3.32}$$

Then v can be expressed as a word of bit-length $O(\log p)$ in $X \cup X^y$ using a calculation similar to (3.31), and the final term in y is a product of permutations from the two copies of A_{p+2}. Another application of Remarks 3.10 and 3.14 completes the proof of (i).

Now (ii) follows from Remark 3.14. □

Remark 3.33. Using Remark 3.10 we see that *every cycle* $(k, k+1, \dots, l)$ *and every element with bounded support in* A_n *has bit-length* $O(\log n)$ *in our generators.*

With a bit more number theory, together with Table 1, we obtain an improvement of Proposition 3.27 that is needed for Theorem C:

Proposition 3.34. *If $n \geq 5$ then S_n and A_n have presentations with* 3 *generators,* 8 *relations and bit-length $O(\log n)$. Moreover, these presentations use a bounded number of exponents, each of which is at most n.*

Proof. We refine the argument in Proposition 3.27. First consider S_n. Here we need to write $n = 2p + 4 - k$ for a prime $p \equiv 2 \pmod 3$ such that $m := p + 2 > k \geq 4$, so that we can use Corollaries 3.12 and 3.24, and then continue as in the proof of Proposition 3.27. In view of the requirements on bit-length and exponents, we also *require that* $p \equiv 11 \pmod{12}$ *if* $n \geq 50$, so that Lemma 3.28 will complete the proof for S_n.

According to Dirichlet's Theorem, for large x there are approximately $x/2 \log x$ primes $\leq x$ of the stated sort, and subtraction yields a prime p in our situation. However, we need a more precise (and effective) result of this type. This is provided in [Mor] (updating [Bre, Er, Mol] with more precise estimates): if $n \geq 50$ then there is such a prime $p \equiv 11 \pmod{12}$. A straightforward examination of the cases $n < 50$ leaves $n = 11, 12$ or 13 to be dealt with. See [Ar, p. 54] for presentations of S_{11}, S_{12} and S_{13} with 2 generators and 6, 7 and 7 relations, respectively (cf. [CoMo, p. 64]). (Note that Table 1 contains the cases $n = 11, 12$, while Corollary 3.12(i) handles the case $n = 13$.)

For A_n we need to write $n = 2p + 4 - k$ for a prime $p \equiv 11 \pmod{12}$ such that $m := p + 2 > k \geq 6$, then use Corollaries 3.8 and 3.24, and again finish as in the proof of Proposition 3.27, using Lemma 3.28. Once again, by [Mol, Mor], if $n \geq 50$ then there is such a prime p. Another straightforward examination leaves the cases $n \leq 13$ and $n = 21, 22, 23, 24, 25, 45, 46, 47, 48$ or 49 to be dealt with. The cases in which $n = p + 2$ or $p + 3$ for some prime p are handled using examples described earlier, and Table 1 handles the remaining cases. (For presentations of A_{11}, A_{12} and A_{13} using 2 generators and 6, 7 and 7 relations, respectively, see [CoMo, p. 67].) □

For the next theorem we will use a presentation in the preceding proposition that is valid for most n. If $n \geq 50$ (or, more precisely, if n is not one of the exceptions mentioned in the above proof) then (3.25) together with one further relation is such a presentation:

$$A_n \text{ or } S_n = \langle X, y \mid R, a = a^y, c = c^y, [d, e^y] = 1, y = w \rangle, \tag{3.35}$$

with $\langle X \mid R \rangle$ in (3.9) or (3.13) for A_n or S_n, respectively, a, c, d, e as in Lemma 3.28, and a suitable word w in $X \cup X^y$ as in (3.30)–(3.32). (The properties required of y and w are described at the end of the proof of Corollary 3.24 and, in gory detail, in the proof of Lemma 3.28.)

We are now able to prove Theorem C:

Theorem 3.36. *If $n \geq 5$ then S_n and A_n have presentations with* 3 *generators,* 7 *relations and bit-length $O(\log n)$. Moreover, these presentations use a bounded number of exponents, each of which is at most n.*

Proof. Let $n = 2m - k$ with $m = p + 2 > k \geq 6$ (cf. the preceding proposition; below we will discuss the existence of a suitable prime p).

Let $G = \langle X \mid R\rangle$ be the presentation in Corollary 3.8(i) or 3.12(ii), so that $|X| = 2$, $|R| = 3$ and one of the following holds:

$$\begin{array}{llll} A_n \text{ case:} & T = \langle \mathbf{a}, \mathbf{b}\rangle = \mathrm{AGL}(1,p)^{(2)} & p \equiv 11 \pmod{12} & G \cong A_m \times T \\ S_n \text{ case:} & T = \langle \mathbf{a}, \mathbf{b}\rangle = \mathrm{AGL}(1,p) & p \equiv 2 \pmod 3 & G/(1\times T) \cong S_m. \end{array}$$

(In the A_n case T has index 2 in $\mathrm{AGL}(1,p)$; in the S_n case G has index 2 in $S_m \times T$.) We also require that $p \equiv 11 \pmod{12}$ in the S_n case when $n \geq 50$.

Let $t \in T$ be such that $T \cap A_m = \langle (t^3)^T\rangle$ (for example, $t = \mathbf{b}^2$ works since $p \equiv 2 \pmod 3$ and the order of $\mathbf{b}$ is odd in the A_n case).

Then the presentation (3.35) of A_n or S_n can be rewritten

$$\langle X, y \mid R, t^3, a^y = a, c^y = c, [d^y, e] = 1, y = w\rangle. \tag{3.37}$$

There was a great deal of freedom in our choice of the elements a, b, c, d, e in the proofs of Lemmas 3.20 and 3.23 (and Corollary 3.24). As in Lemma 3.28, we now *choose* a so that its image in the alternating or symmetric group $G/(1 \times T)$ is a 3-cycle (for the S_n case, this requires c to be an odd permutation and $l \geq 3$ in the proof of Lemma 3.15).

The presentation (3.37) has 8 relations and bit-length $O(\log n)$.

We use the following additional ingredients:

- Write $a \in G$ as $a = (a_1, *) \in A_m \times T$ with a_1 of order 3.
- Let $\bar{a}$ and $\hat{a}$ be words in X such that $\bar{a} = (a_1, 1)$ and $\hat{a} = (a_1, t)$ when evaluated in G.

We claim that the 7*-relator group*

$$J := \langle X, y \mid R, \bar{a}^y = \hat{a}, c^y = c, [d^y, e] = 1, y = w\rangle \tag{3.38}$$

is isomorphic to the group in (3.37).

For, we can view $G = \langle X\rangle \leq J$. Then $a, \bar{a}, \hat{a}, c, d, e \in J$.

Since both $\bar{a}^y = \hat{a} = (a_1, t)$ and a_1 have order 3 we have $(1, t)^3 = 1$ in J. Thus, J satisfies the presentation (3.37) and hence is as claimed.

Bit-length: As in the proof of Theorem 3.27, a, c, d and e can be expressed as words in X of bit-length $O(\log n)$ mod T. Since $T = \langle \mathbf{a}\rangle\langle \mathbf{b}\rangle$ in Corollaries 3.8 and 3.12, $\bar{a}$ and $\hat{a}$ have bit-length $O(\log n)$ as well.

Finally, we need to discuss whether we have handled all groups A_n and S_n; or, what amounts to the same thing, for which n a prime p can be found satisfying all of the conditions we have imposed.

As in Proposition 3.34, (3.38) takes care of A_n except for the cases $n \leq 13$ and $n = 21, 22, 23, 24, 25, 45, 46, 47, 48, 49$; and these are handled exactly as in that proposition.

For the S_n case we have imposed a further condition beyond what was used in Proposition 3.34: we need to write $n = 2p + 4 - k$ for a prime $p \equiv 2 \pmod 3$ such that $m = p + 2 > k \geq l + 2 \geq 5$, and $p \equiv 11 \pmod{12}$ if $n \geq 50$. (The conditions in Corollary 3.24 were $m = p + 2 > k \geq l + 2 \geq 4$, but here we need to be able to find a 3-cycle a in A_l.) Once again these requirements can be met for all n except if $n < 6$ or $n = 9, 10, 11$, and those cases can be handled as before. □

None of the presentations in this or the preceding section has length $O(\log n)$.

By Remarks 3.10 and 3.14, the exponents in (3.31) are all less than n; there is also an exponent $p + 1 - k$ used to write u in that remark. As already noted, these presentations have *bounded expo-length* (cf. Section 2). See Remark 8 in Section 11 for comments concerning the boundedness of expo-length for other families of almost simple groups.

Before continuing, we note the following simple improvement of [GKKL1, Lemma 2.1].

Lemma 3.39. *Let $G = \langle D\rangle$ be a finite group having a presentation $\langle X \mid R\rangle$; let $\pi\colon F_X \to G$ be the natural map from the free group F_X on X. Then G also has a presentation $\langle D \mid R'\rangle$ such that $|R'| = |D| + |R| - |\pi(X) \cap D|$.*

Proof. We recall the simple idea used in the proof of [GKKL1, Lemma 2.1]. Write each $x \in X$ as a word $v_x(D)$ in D, and each $d \in D$ as a word $w_d(X)$ in X; and let $V(D) = \{v_x(D) \mid x \in X\}$. According to the proof of [GKKL1, Lemma 2.1], we then obtain another presentation for G:

$$G = \langle D \mid d = w_d(V(D)),\, r(V(D)) = 1, d \in D, r \in R\rangle.$$

For each $d \in \pi(X) \cap D$, one of the above relations can be taken to be $d = d$, and hence can be deleted. □

In view of the desire in [BCLO] for specific generators (namely, $(1, 2)$ and $(1, 2, \dots, n)$), we note the following consequence of the preceding theorem and lemma:

Corollary 3.40. *Let $G = A_n$ or S_n, $n \geq 5$.*

(i) *If a and b are any generators of G, then there is a presentation of G using 2 generators that map onto a and b, and 9 relations.*
(ii) *There is a presentation of G using 2 generators and 8 relations.*

We do not have information concerning the bit-length of any of the resulting presentations.

Proof. Part (i) follows from Theorem 3.36 and the preceding lemma, using $|\pi(X) \cap D| \geq 0$.

For (ii), note that we have provided a presentation $\langle X \mid R\rangle$ for G such that some element of X projects onto an element $a \in G$ that is either a 3-cycle (z in Lemma 3.3) or has a power that is a 3-cycle (such as g in Corollary 3.8(ii) or Proposition 3.18). Let b be any element of G such that $G = \langle a, b\rangle$. Now use $D = \{a, b\}$ in the preceding corollary (compare Section 11, Remark 4). □

3.5. **An explicit presentation for S_n.**

The presentations in Sections 3.1 and 3.2 are not difficult to understand, and they visibly encode information concerning various alternating and symmetric groups. However, the presentations in Theorem 3.36 are not as explicit as one might wish. Therefore, we will provide a presentation of S_n for n odd (see Remark 3.41 for even n). Although this presentation is in no sense elegant or informative, it does have the significant advantage of using only 7 relations.

Find a prime $p \equiv 11 \pmod{12}$ such that $n - 1 \geq p \geq (n + 2)/2$. (This places a mild restriction on n, as seen in the proof of Theorem 3.36 . For $n \geq 50$ there is always such a prime.)

Let $k = 2p + 4 - n$, so that $p + 2 > k \geq 6$. Then $k \equiv n \equiv 1 \pmod 2$, so that $p - k$ is even.

The desired presentation is

$$S_n = \langle \mathbf{a}, \mathbf{g}, \mathbf{y} \mid \mathbf{a}^p = (\mathbf{g}^3)^{p-1}, (\mathbf{a}^s)^{\mathbf{g}^3} = \mathbf{a}^{s-1}, \big(\mathbf{g}^{p-1}(\mathbf{g}^{p-1})^{\mathbf{a}}\big)^2 = 1,$$
$$a^{\mathbf{y}} = \hat{a}, c^{\mathbf{y}} = c, [d^{\mathbf{y}}, e] = 1, \mathbf{y} = w\rangle,$$

for words $a, c, d, e, \hat{a}, w$ defined below and integers r and s such that $s(r-1) \equiv -1 \pmod{p}$ and $\mathbb{F}_p^* = \langle r\rangle$.

Notes: In order to conform with the notation in Section 3.1, we view $\mathrm{AGL}(1,p)$ as acting on $\{1,\dots,p\}$ with $\mathbf{a} \equiv (1,\dots,p) \in \mathrm{AGL}(1,p)$, but we use p in place of 1 in Lemma 3.3. Finally, we use $\mathbf{a}$ in place of a since the latter plays a prominent role in Section 3.3.

The permutations in S_n indicated below are not part of the presentation, but are provided in order to help keep track of the map $\langle \mathbf{a}, \mathbf{g}, \mathbf{y}\rangle \to S_n$ into the symmetric group on the n points (3.29). The notation used here should **not** be viewed mod p.

(1) $\mathbf{z} := \mathbf{g}^{p-1} \equiv (p, p+1, p+2)$,
$\mathbf{b} := \mathbf{g}^3$ (so that $\langle \mathbf{a}, \mathbf{b}\rangle = \mathrm{AGL}(1,p)$),

(2) $\mathbf{z}(i) := \mathbf{z}^{\mathbf{a}^{-i}} \equiv (i, p, p+1)$ for $1 \le i \le p$,
$c(i,j) := \mathbf{z}(-i)\mathbf{z}(-j)^{-1}\mathbf{z}(-i) \equiv (i,j)(p+1, p+2)$ for $1 \le i < j \le p$,
$c_i := \big(c(1,2)\mathbf{a}\big)^{i-2} c(1,2)\mathbf{a}^{-(i-2)} \equiv (1, 2, \dots, i)$ with i odd and $3 \le i \le p$ (cf. (3.15)).

(3) (Constructing a transposition)
$b_2 := \mathbf{b}^{(p-1)/2} \equiv (1, p-1)(2, p-2) \cdots \big((p-1)/2, p-(p-1)/2\big)$,
$c_\bullet := c_{(p-1)/2} c_{(p-1)/2}^{-\mathbf{a}^{(p+1)/2}} \equiv \big(1, 2, \dots, (p-1)/2\big)\big(p-1, p-2, \dots, p-(p-1)/2\big)$,
$v := \big(c(1,p-1)c_\bullet^{-1}\big)^{(p-1)/2} c(1,p-1) c_\bullet^{(p-1)/2}$
$\equiv (1, p-1)(2, p-2) \cdots \big((p-1)/2, p-(p-1)/2\big)(p+1, p+2)$ (cf. (3.15)),
$t := v b_2 c(1,2) \equiv (1,2)$.

(4) $a := \mathbf{z}(3)^{\mathbf{z}(1)\mathbf{z}(2)} \equiv (1,2,3)$,
$c := t c_k \equiv (2, \dots, k)$ (an odd permutation),
$d := c_3^{-1}\mathbf{a}\mathbf{z} \equiv (3, \dots, p+2)$,
$e := \mathbf{a} c_k^{-1} t \mathbf{z} \equiv (1, 2, k+1, \dots, p+2)$, so that $e^{\mathbf{y}} \equiv (1, 2, 0, -1, \dots, -p+k-1)$ (also odd permutations).

(5) $\hat{a} := a\big(\mathbf{b}^2 z(1) z(-1)\big)^{(p+1)/2} \in \{a\} \times T$.

(6) $x := [c(1,k+2)c(2,k+1)]^{c(1,k+2)^{\mathbf{y}} c(2,k+1)^{\mathbf{y}}}$
$\equiv (-1, k+2)(0, k+1) = [(1, k+2)(2, k+1)]^{(-1,1)(0,2)}$,
$u^{-1} := (\mathbf{a}\mathbf{z})(\mathbf{a}\mathbf{z})^{\mathbf{y}}$
$\equiv (1, \dots, k, k+1, k+2, \dots, p, p+1, p+2)(1, \dots, k, 0, -1, \dots, -p+k-1)$,
$w := (xu^2)^{(p-k)/2} x u^{-(p-k)}$
$\equiv (-p-1+k, p+2)(-p+k, p+1) \cdots (-1, k+2)(0, k+1)$ (cf. (3.15).

Remarks 3.41. 1. In the isomorphism given by (3.4), $\mathbf{z}$ maps to an element of $A_{p+2} \times \{1\}$. Hence, the element a defined above also maps into $A_{p+2} \times \{1\}$, so that the element $\bar{a}$ used in (3.38) is just our a. The remainder of the presentation given above is just a straightforward translation from Section 3.4.

2. A presentation for the alternating groups is similar but slightly simpler: only even permutations are involved.

3. *Changes needed when n and k are even*:

$(4')$ $c := c_{k-1}t \equiv (1, \dots, k)$ (an odd permutation),
$e := \mathbf{a}tc_{k-1}^{-1}\mathbf{z} \equiv (1, k+1, \dots, p+2)$ (another odd permutation).

$(6')$ $w := t^{\mathbf{az}}t^{\mathbf{az}y}t^{\mathbf{az}}(xu^2)^{(p-k-1)/2}xu^{-(p-k-1)}$
$\equiv (-p-1+k, p+2)(-p+k, p+1)\cdots(-1, k+2)(0, k+1)$ as before.

3.6. **Weyl groups.** It is easy to use Theorem 3.36 to obtain presentations for the Weyl groups of types B_n or D_n. However, we leave this to the reader, dealing instead with a subgroup W_n of those Weyl groups that is needed later (in Section 10).

Let $W_n := \mathbb{Z}_2^{n-1} \rtimes A_n$ be the subgroup of the monomial group of $\mathbb{R}^n$ such that $\mathbb{Z}_2^{n-1}$ consists of all ± 1 diagonal matrices of determinant 1, and the alternating group A_n permutes the standard basis vectors of $\mathbb{R}^n$.

Proposition 3.42. *If $n \geq 4$ then W_n has a presentation with 4 generators, 11 relations and bit-length $O(\log n)$. If $n = 4$ or 5 then W_n also has a presentation with 3 generators and 7 relations.*

Proof. We first consider the case $n \geq 5$. Let $\langle X \mid R \rangle$ be a presentation for $A = A_n$. Let $\sigma = (1, 2, 3) \in A$, choose a 2-element generating set of $H := A_{\{1,2\}}$, and consider the group J with the following presentation:

Generators: X, s (where s represents $\operatorname{diag}(-1, -1, 1, \dots, 1)$).

Relations:
(1) R.
(2) $s^2 = 1$.
(3) $[s, H] = 1$.
(4) $ss^{\sigma}s^{\sigma^2} = 1$.

There is an obvious surjection $\pi\colon J \to W_n$. We can view $A = \langle X \rangle \leq J$. By (3), $\binom{n}{2} \geq |s^A| \geq |\pi(s^A)| = \binom{n}{2}$, so that s^A can be identified with the 2-sets in $I = \{1, \dots, n\}$. Thus, there are well-defined elements $s_{ij} = s_{ji} \in s^A$ for all distinct $i, j \in I$.

By (4), $s_{1j}s_{jk}s_{k1} = 1$ whenever $1, j, k$ are distinct. Since all s_{ij} are involutions, it follows that $s = s_{12}$ commutes with all s_{1j}, and hence also with all s_{jk}, so that $N := \langle s^A \rangle$ is elementary abelian. Then $J = AN$ has order $|W_n|$.

Now Theorem 3.36 yields the stated numbers of generators and relations for $n \geq 4$. For $n = 4$ or 5 we instead use (3.2). □

4. Rank 1 groups

4.1. **Steinberg presentation.** Each rank 1 group G we consider has a Borel subgroup $B = U \rtimes \langle h \rangle$, with U a p-group. There is an involution t (mod $Z(G)$ in the case $\mathrm{SL}(2, q)$ with q odd) such that $h^t = h^{-1}$ (or h^{-q} in the unitary case). The *Steinberg presentation* for these groups [St2, Sec. 4] consists of

- a presentation for B,
- a presentation for $\langle h, t \rangle$, and
- $|U| - 1$ relations of the form

$$u_0^t = u_1 h_0 t u_2, \tag{4.1}$$

with u_0, u_1, u_2 nontrivial elements of U and $h_0 \in \langle h \rangle$ (one relation for each choice of u_0).

4.2. **Polynomial notation.** Our groups will always come equipped with various elements having names such as u or h. For any polynomial $g(x) = \sum_0^e g_i x^i \in \mathbb{Z}[x]$, $0 \le g_i < p$, define powers as follows:

$$[[u^{g(x)}]]_h = (u^{g_0})(u^{g_1})^{h^1} \cdots (u^{g_e})^{h^e}, \tag{4.2}$$

so that

$$[[u^{g(x)}]]_h = u^{g_0} h^{-1} u^{g_1} h^{-1} u^{g_2} \cdots h^{-1} u^{g_e} h^e \tag{4.3}$$

by "Horner's Rule" [GKKL1, (4.14)] (compare (3.15)).

As in [GKKL1, Sec. 4.3], we need to be careful about rearranging the terms in (4.2) when not all of the indicated conjugates of u commute.

4.3. $\mathbf{SL(2, \mathit{q})}$. In [CRW] there is a presentation for $\mathrm{PSL}(2,q)$ with at most 13 relations; and it follows readily from that presentation that $\mathrm{SL}(2,q)$ has one with at most 17 relations. We now provide presentations having fewer relations, based on the matrices

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} \zeta^{-1} & 0 \\ 0 & \zeta \end{pmatrix}. \tag{4.4}$$

Theorem 4.5. $\mathrm{SL}(2,q)$ *and* $\mathrm{PSL}(2,q)$ *have presentations with* 3 *generators*, 9 *relations and bit-length* $O(\log q)$. *When q is even* $\mathrm{PSL}(2,q)$ *has a presentation with* 3 *generators*, 5 *relations and bit-length* $O(\log q)$.

Proof. By [CoMo, pp. 137-138], if $q \le 9$ then $\mathrm{SL}(2,q)$ and $\mathrm{PSL}(2,q)$ have presentations with 2 generators and at most 4 relations. Assume that $q > 9$, let ζ be a generator of $\mathbb{F}_q^*$, and let $k, l \in \mathbb{Z}$ be such that $\zeta^{2k} = \zeta^{2l} + 1$ and $\mathbb{F}_q = \mathbb{F}_p[\zeta^{2k}]$ (as in [GKKL1, Section 3.5.1]).

Set $d = \gcd(k,l)$. Then $\mathbb{F}_q = \mathbb{F}_p[\zeta^{2d}]$.

Let $m(x) \in \mathbb{F}_p[x]$ be the minimal polynomial of ζ^{2d}. If $\gamma \in \mathbb{F}_q$, let $g_\gamma(x) \in \mathbb{F}_p[x]$ satisfy $g_\gamma(\zeta^{2d}) = \gamma$ and $\deg g_\gamma < \deg m$.

We will show that $G = \mathrm{SL}(2,q)$ is isomorphic to the group J having the following presentation.

Generators: u, t, h.

Relations:

(1) $u^p = 1$.
(2) $u^{h^k} = u u^{h^l} = u^{h^l} u$.
(3) $[[u^{m(x)}]]_{h^d} = 1$ in the notation of (4.2).
(4) $u^h = [[u^{g_{\zeta^2}(x)}]]_{h^d}$.
(5) $[t^2, u] = 1$ (or $t^2 = 1$ in the case $\mathrm{PSL}(2,q)$ with q odd).
(6) $h^t = h^{-1}$.
(7) $t = u u^t u$.
(8) $ht = [[u^{g_{\zeta^{-1}}(x)}]]_{h^d} \, [[u^{g_\zeta(x)}]]^t_{h^d} \, [[u^{g_{\zeta^{-1}}(x)}]]_{h^d}$.

Matrix calculations using (4.4) easily show that there is a surjection $J \to G$. By (1), (2) and [GKKL1, Lemma 4.1] (compare [Bau, CR1, CRW]), $U := \langle u^{\langle h^k, h^l \rangle} \rangle$ is elementary abelian; since $d = \gcd(k,l)$ we have $U = \langle u^{\langle h^d \rangle} \rangle$. By (1) and (3) we can identify U with the additive group of $\mathbb{F}_q$ in such a way that h^d acts as multiplication by ζ^{2d}. By (4), h acts on U as a transformation of order $(q-1)/(2, q-1)$. (Note that U is defined using h^k and h^l rather than h so that [GKKL1, Lemma 4.1] can be used.)

By (7) and (8), $J = \langle U, U^t\rangle$, and J is perfect since $U = [U, h]$. Moreover, using (6) we see that $z := h^{(q-1)/(2,q-1)}$ is inverted by t, centralizes U and U^t, and hence is an element of $Z(J)$ having order 1 or 2.

Thus, $\langle u, h\rangle/\langle z\rangle$ is isomorphic to a Borel subgroup of $\mathrm{PSL}(2, q)$. By (6), $\langle h, t\rangle/\langle z\rangle$ is dihedral of order $2(q-1)/(2, q-1)$.

We already know that $\langle h\rangle$ acts on the nontrivial elements of U with at most 2 orbits, with orbit representatives u^1 and $[[u^{g_\zeta(x)}]]_{h^d}$ if q is odd. As in the proof of [GKKL1, Sec. 4.4.1], (7) and (8) provide the relations (4.1) required to let us deduce that $J/\langle z\rangle \cong \mathrm{PSL}(2, q)$.

Now J is a perfect central extension of $\mathrm{PSL}(2, q)$, and hence is $\mathrm{SL}(2, q)$ or $\mathrm{PSL}(2, q)$. Finally, (5) distinguishes between these groups when q is odd. The bit-length of the presentation is clear from (4.3).

Finally, if q is even there are significant simplifications. We may assume that $k = d = 1$, so that $h^d = h$ acts on $U = \langle u^{\langle h^k, h^l\rangle}\rangle$; the induced automorphism has order $q-1$ by (3), and (4) can be deleted. Relation (5) can be deleted since (1) and (7) imply that $t^2 = 1$; and (8) is not needed since $\langle h\rangle$ has only one orbit on the nontrivial elements of U. □

Remark 4.6. *Every element of* $\mathrm{SL}(2, q)$ *has bit-length* $O(\log q)$ *in our generators.* For, this is true of all elements of U by (4.3), while $\mathrm{SL}(2, q) = UU^tU$.

Remark 4.7. If $d = 1$ then relation (4) can be removed since then $h^d = h$ acts as multiplication by ζ^2, by (3).

In [GKKL1, Section 3.5.1] it was observed that we can choose $k = 1$, $l = 1$ or $k = 2$. Thus, $d \le 2$ for some choice of k and l. If q is even then $d = 1$. If $q \equiv 3 \pmod 4$ and $k = 2$ we can change ζ to $-\zeta^2$ in order to obtain $k = 1$ and hence $d = 1$. We can also prove that there are choices for ζ, k, l that yield $d = 1$ when $q \equiv 5 \pmod 8$, but we do not know how to obtain such choices in general.

Remark 4.8. Now that we have the notation in (4.2), we can give more examples along the lines of Examples 3.19(1) and (3). Let q be a power of an odd prime p such that $(3, q-1) = 1$; we may assume that $q > 5$.

(1′) *S_{q+2} has a presentation with 2 generators and 6 relations.* For, if ζ is as above, $\zeta + 1 = \zeta^s$, and $g(x)$ is the minimal polynomial of ζ over $\mathbb{F}_p$, then $\mathrm{AGL}(1, q) \cong \langle u, h \mid u^p = h^{q-1}, u^{h^s} = uu^h = u^hu, [[u^{g(x)}]]_h = 1\rangle$; this is proved as above, using [GKKL1, Lemma 4.1]. Then Proposition 3.18 provides the stated presentation.

(3′) *A_{q+2} has a presentation with 2 generators and 6 relations – only 5 relations if $q \equiv 3 \pmod 4$.* This time let k, l and $m(x)$ be as in the proof of the preceding theorem, and obtain $\mathrm{AGL}(1, q)^{(2)} \cong \langle u, h \mid u^p = h^{(q-1)/2}, u^{h^k} = uu^{h^l} = u^{h^l}u, [[u^{m(x)}]]_h = 1\rangle$, after which Proposition 3.18 provides the stated presentation (with $\rho = (2, (q-1)/2)$).

4.4. **Unitary groups.** We will obtain presentations for 3-dimensional unitary groups by taking the presentations in [GKKL1, Sec. 4.4.2] and deleting the portions that were needed to produce short presentations. We use matrices of the form

$$(4.9)\quad u = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & -\bar{\alpha} \\ 0 & 0 & 1 \end{pmatrix}, w = \begin{pmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, t = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, h = \begin{pmatrix} \bar{\zeta}^{-1} & 0 & 0 \\ 0 & \bar{\zeta}/\zeta & 0 \\ 0 & 0 & \zeta \end{pmatrix}$$

with $\alpha, \beta, \gamma \in \mathbb{F}_{q^2}$ arbitrary such that $\beta + \bar{\beta} = -\alpha\bar{\alpha} \neq 0$ and $\gamma = -\bar{\gamma} \neq 0$.

Theorem 4.10. $\mathrm{SU}(3, q)$ *and* $\mathrm{PSU}(3, q)$ *have presentations with* 3 *generators*, 23 *and* 24 *relations*, *respectively*, *and bit-length* $O(\log q)$.

Proof. Let ζ be a generator of $\mathbb{F}_{q^2}^*$. As in [GKKL1, Sec. 4.4.2], we will assume that $q \neq 2, 3, 5$ and use elements $a = \zeta^k, b = \zeta^l$, where $1 \leq k, l < q^2$, such that

$$a^{2q-1} + b^{2q-1} = 1 \;\text{ and }\; a^{q+1} + b^{q+1} = 1,$$
$$\mathbb{F}_q = \mathbb{F}_p[a^{q+1}], \text{ and}$$
$$\mathbb{F}_{q^2} = \mathbb{F}_p[a^{2q-1}] \text{ if } q \text{ is odd while } \mathbb{F}_q = \mathbb{F}_p[a^{2q-1}] \text{ if } q \text{ is even.}$$

Let $d = \gcd(k, l)$.

If $\gamma \in \mathbb{F}_{q^2}$ write $\gamma' := \gamma^{q+1}$ and $\gamma'' := \gamma^{2q-1}$, and also let $m_\gamma(x)$ denote its minimal polynomial over $\mathbb{F}_p$. If $\delta \in \mathbb{F}_p[\gamma]$, let $f_{\delta;\gamma}(x) \in \mathbb{F}_p[x]$ with $f_{\delta;\gamma}(\gamma) = \delta$ and $\deg f_{\delta;\gamma} < \deg m_\gamma$ (compare [GKKL1, Sec. 4.4.2]).

The required presentation is as follows:

Generators: u, h, t.

Relations:

(1) $w = w^{h^k} w^{h^l} = w^{h^l} w^{h^k}$, where w is defined by $u = u^{h^k} u^{h^l} w$.
(2) $w^p = 1$.
(3) $[[w^{m_{a'}(x)}]]_{h^d} = 1$.
(4) $[[w^{f_{\zeta';a'}(x)}]]_{h^d} = w^h$.
(5) $u = u^{h^l} u^{h^k} w_1$.
(6) $[u, w] = [u^{h^k}, w] = 1$.
(7) $u^p = w_2$.
(8) $[[u^{m_{a''}(x)}]]_{h^d} = w_3$.
(9′) $[[u^{f_{\zeta'';a''}(x)}]]_{h^d} = uw_4$ if q is odd.
(9″) $([[u^{f_{\alpha;a''}(x)}]]_{h^d})^h [[u^{f_{\beta;a''}(x)}]]_{h^d} = u^{h^2} w_5$ if q is even and ζ'' satisfies $\zeta''^2 = \alpha\zeta'' + \beta$ for $\alpha, \beta \in \mathbb{F}_q$.
(10) $[u, u^h] = w_6$ and $[u^{h^k}, u^h] = w_7$ if q is even.
(11) $t^2 = 1$.
(12) $h^t = h^{-q}$.
(13) $u_i^t = u_{i1} h_i t u_{i2}$ for $1 \leq i \leq 7$, relations due to Hulpke and Seress [HS].
(14) $1 = v_{11} v_{12}^t v_{13} t$ and $h = v_{21} v_{22}^t v_{23} t$.

Here, the elements w_i are specific words in of $w^{\langle h^d \rangle}$, the elements u_{ij}, v_{ij} are specific words in $u^{\langle h^d \rangle}$, and the elements h_i are specific powers of h; all depend on the initial choice of u and ζ.

Note that $\langle w^{\langle h^d \rangle} \rangle$ is defined using h^d rather than h so that [GKKL1, Sec. 4.1] can be used together with (1) and (5). See [GKKL1, Sec. 4.4.2] for a proof that this is, indeed, a presentation of $\mathrm{SU}(3, q)$. As in [GKKL1, Sec. 4.4.2], one further relation of bit-length $O(\log q)$ produces a presentation for $\mathrm{PSU}(3, q)$. □

Remark 4.11. Let $U := \langle u^{\langle h^d \rangle} \rangle$. By applying (4.3) to both $Z(U) = \langle w^{\langle h^d \rangle} \rangle$ and $U/Z(U)$, we find that all elements of U have bit-length $O(\log q)$, and hence the same holds for $\mathrm{SU}(3, q) = UU^tUU^tU$.

4.5. **Suzuki groups.** The short presentation in [GKKL1, Section 4.4.3] uses 7 generators and 43 relations for $\mathrm{Sz}(q)$. This can be used for ${}^2F_4(q)$ in Section 7.

Nevertheless, we note that there is an easy modification similar to what occurred for $\mathrm{SL}(2,q)$ and $\mathrm{SU}(3,q)$: three generators are powers of a fourth, allowing us to decrease to 4 generators and 31 relations.

5. $\mathrm{SL}(3,q)$

The groups $\mathrm{SL}(3,q)$ will reappear more often in the rest of the proof than any other rank 2 groups: we will use $\mathrm{SL}(3,q)$ to obtain first all higher-dimensional groups $\mathrm{SL}(n,q)$, and then all higher-dimensional classical groups. Therefore we will be more explicit with these groups than the other rank 2 groups (cf. Section 7).

Theorem 5.1. *$\mathrm{SL}(3,q)$ has a presentation with 4 generators, 14 relations and bit-length $O(\log q)$.*

Proof. When $q \le 9$ there are presentations with 2 generators and at most 10 relations [CMY, CR3]. (These cases also can be handled directly by a slight variation on the following approach; compare the end of Section 7.1.) Hence, we will assume that $q > 9$.

Let $\mathrm{SL}(2,q) \cong L = \langle X \mid R\rangle$, with $X = \{u, t, h\}$ and $\langle \zeta\rangle = \mathbb{F}_q^*$ as in the proof of Theorem 4.5. We view the elements of $\mathrm{SL}(3,q)$ as matrices, with L consisting of the matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & 1\end{smallmatrix}\right)$.

We will show that $G = \mathrm{SL}(3,q)$ is isomorphic to the group J having the following presentation.

Generators: X and c (corresponding to the permutation matrix $\left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0\end{smallmatrix}\right)$ acting as $(3,2,1)$).

Relations:

(1) R.

(2) $c^t = t^2c^2$.

(3) $hh^ch^{c^2} = 1$.

(4) $u^{h^c} = u^{\mathrm{diag}(1,\zeta^{-1})}$, written as a word in X. (The matrix $\mathrm{diag}(1,\zeta^{-1})$ is not in L if q is odd, but it can be viewed as inducing an automorphism of L.)

(5) $[u^c, u] = (u^t)^{c^2}$.

(6) $[u^c, u^t] = 1$.

Matrix calculations easily show that there is a surjection $J \to G$. There is a subgroup of J we identify with $L = \langle X\rangle$. We separate the argument into four steps.

1. *Computations in $\langle t, c\rangle$.* The relations $t^4 = 1$ and (2) imply that

$$tct = c^2, \quad t^{-1}c^2t^{-1} = c, \quad t^{-1}c^3t = t^{-1}c^2t^{-1}tct = cc^2, \tag{5.2}$$

and hence

$$t^c = c^{-1}tctt^{-1} = c^{-1}c^2t^{-1} = ct^{-1},\ t^{c^2} = (ct^{-1})^c = t^{-1}c. \tag{5.3}$$

It follows that

$$t^2(t^2)^{c^2}(t^2)^c = t^2 \cdot t^{-1}ct^{-1}c \cdot ct^{-1}ct^{-1} = tc(t^{-1}c^2t^{-1})ct^{-1} = tccct^{-1} = c^3. \tag{5.4}$$

2. *The action of h, h^c, h^{c^2}, c^3 on L.* By (5.2) and (3), $(h^c)^t = h^{t^{-1}c^2} = (h^{-1})^{c^2} = hh^c$. Consequently, conjugating both sides of (4) by t gives

$$\left(u^{h^c}\right)^t = \left(u^t\right)^{h^{ct}} = \left(u^t\right)^{hh^c} = \left(u^{th}\right)^{h^c}$$
$$\left(u^{\mathrm{diag}(1,\zeta^{-1})}\right)^t = \left(u^t\right)^{\mathrm{diag}(1,\zeta^{-1})^t} = \left(u^t\right)^{\mathrm{diag}(\zeta^{-1},1)} = \left(u^t\right)^{h\,\mathrm{diag}(1,\zeta^{-1})} = \left(u^{th}\right)^{\mathrm{diag}(1,\zeta^{-1})}.$$

(Recall that $h = \mathrm{diag}(\zeta^{-1}, \zeta)$ in Section 4.3.) Thus, by (4), h^c acts on $L = \langle u, u^{th}\rangle$ as conjugation by $\mathrm{diag}(1, \zeta^{-1})$. We know how h acts on L since $h \in L$. Now (3) implies that h^{c^2} acts on L as conjugation by $\mathrm{diag}(\zeta^{-1}, 1)$.

If q is odd then $t^2 = h^{(q-1)/2}$ is in $Z(L)$. It follows that t^2, $(t^2)^c = (h^c)^{(q-1)/2}$ and $(t^2)^{c^2} = (h^{c^2})^{(q-1)/2}$ act on L as they should: as conjugation by $1, \mathrm{diag}(1, -1)$ and $\mathrm{diag}(1, -1)$, respectively. Now (5.4) implies that $c^3 = t^2(t^2)^{c^2}(t^2)^c$ acts trivially on $L = \langle X\rangle$ and hence on $J = \langle X, c\rangle$. This also holds trivially if q is even.

3. *The elements* $e_{ij}(\lambda)$. For all integers m and all $\lambda \in \mathbb{F}_q$, write

$$\begin{aligned}
&e_{12}(\zeta^m) := u^{(h^m)^c},\ e_{12}(0) := 1 && e_{21}(\lambda) := e_{12}(-\lambda)^t\\
&e_{23}(\lambda) := e_{12}(\lambda)^c && e_{32}(\lambda) := e_{21}(\lambda)^c\\
&e_{31}(\lambda) := e_{12}(\lambda)^{c^2} && e_{13}(\lambda) := e_{21}(\lambda)^{c^2}.
\end{aligned}$$

Then $e_{12}(1) = u$, $e_{23}(1) = u^c$, and $e_{12}(\mathbb{F}_q)$ is an elementary abelian subgroup of L by (4). Then we also have $e_{21}(\mathbb{F}_q) < L$. By (5.3), $c = tt^c \in \langle X \cup X^c\rangle$, so that J is generated by the elements $e_{ij}(\lambda)$.

Clearly, $\langle c\rangle$ acts on the set of subgroups $e_{ij}(\mathbb{F}_q)$; in fact $\langle t, c\rangle$ acts as the symmetric group S_3 on subscripts (this is the Weyl group of G). For example, by (5.2), $e_{23}(\mathbb{F}_q)^t = e_{12}(\mathbb{F}_q)^{ct} = e_{12}(\mathbb{F}_q)^{t^{-1}c^2} = e_{13}(\mathbb{F}_q)$; and (as we have seen) t^2 acts correctly on $L^{c^{-1}} = L^{c^2}$ and hence also on $e_{12}(\mathbb{F}_q)^{c^2} = e_{31}(\mathbb{F}_q)$. Similarly, t acts correctly on each $e_{ij}(\mathbb{F}_q)$.

4. *Verification of the Steinberg relations* (see [GKKL1, Section 5.1 or 5.2]). The relations

$$e_{ij}(\lambda)e_{ij}(\mu) = e_{ij}(\lambda + \mu)$$

follow from the corresponding relation in L (with $\{i, j\} = \{1, 2\}$) by conjugating with t and c. We will deduce the remaining relations from (5) and (6) by conjugating with t, c, h, h^c and h^{c^2}; we have seen that these act on the set of subgroups $e_{ij}(\mathbb{F}_q)$ as they do in G.

We have $[e_{23}(1), e_{21}(1)] = [u^c, u^t] = 1$ by (6). Conjugating by $h^x(h^c)^y$ we obtain

$$[e_{23}(\zeta^{2y-x}), e_{21}(\zeta^{-2x+y})] = 1.$$

This does not cover all relations of the form $[e_{23}(\lambda), e_{21}(\mu)] = 1$ since $\det\left(\begin{smallmatrix}-1 & 2\\ -2 & 1\end{smallmatrix}\right) = 3$, but this does imply that

$$[e_{23}(1), e_{21}(\mu^3)] = 1 \quad \text{for all } \mu \in \mathbb{F}_q.$$

The additive subgroup of $\mathbb{F}_q$ generated by the cubes in $\mathbb{F}_q^*$ is closed under multiplication, and so is a subfield of size $\geq 1 + (q-1)/3$ and hence is all of $\mathbb{F}_q$ since $q \neq 4$. It follows that

$$[e_{23}(1), e_{21}(\mu)] = 1 \quad \text{for all } \mu \in \mathbb{F}_q.$$

Conjugating this by all h^x yields all relations of the form

$$[e_{23}(\lambda), e_{21}(\mu)] = 1 \quad \text{for all } \lambda, \mu \in \mathbb{F}_q.$$

Conjugating by $\langle t, c\rangle$, we obtain

$$(5.5)\quad [e_{kl}(\lambda), e_{km}(\mu)] = [e_{km}(\lambda), e_{lm}(\mu)] = 1 \quad \text{for all distinct } k, l, m \text{ and all } \lambda, \mu.$$

Similarly, (5) implies that

$$[e_{23}(1), e_{12}(1)] = [u^c, u] = (u^t)^{c^2} = e_{13}(-1).$$

Since t acts correctly on each $e_{ij}(\mathbb{F}_q)$, conjugating by t gives $[e_{13}(-1), e_{21}(1)] = e_{23}(1)$. Conjugating by $h^x(h^c)^y$ we obtain

$$[e_{23}(\zeta^{2y-x}), e_{12}(\zeta^{2x-y})] = e_{13}(-\zeta^{x+y}), \quad [e_{13}(\zeta^{x+y}), e_{32}(\zeta^{x-2y})] = e_{12}(\zeta^{2x-y}).$$

Once again these do not cover all relations of the form

$$[e_{23}(\lambda), e_{12}(\mu)] = e_{13}(-\lambda\mu) \quad \text{and} \quad [e_{13}(\lambda), e_{32}(\mu)] = e_{12}(\lambda\mu).$$

However, using the standard identity $[x, ab] = [x, b][x, a]^b$, together with (5.5) and the fact that the cubes in $\mathbb{F}_q^*$ generate $\mathbb{F}_q$ under addition, we deduce all such relations.

Conjugating by $\langle t, c\rangle$ yields all remaining Steinberg relations. Then J is a homomorphic image of $G = \mathrm{SL}(3, q)$, and hence $J \cong G$. By Theorem 4.5 and Remark 4.6, our presentation has the required bit-length; note that c is the product of an element of L and an element of L^c. The presentation uses $|R| + 5 = 14$ relations. □

In order to obtain a presentation for $\mathrm{PSL}(3, q)$ when $m := (q-1)/3$ is an integer, add the relation $h^m(h^{2m})^c = 1$.

Note that the computations in the above proof would be considerably simpler if we added the relation $c^3 = 1$; this is a relation we will have available in the proof of Theorem 6.1.

Remark 5.6. Using Remark 4.6 we see that *each element of* $\mathrm{SL}(3, q)$ *has bit-length* $O(\log q)$ *in our generators.*

Remark 5.7. For future use we will need to know that $\mathrm{SL}(3, q) = \langle a, a^{(1,2,3)}\rangle$ *for some* $a \in L$ *of order* $q + 1$. For, $\langle a, a^{(1,2,3)}\rangle$ is an irreducible subgroup of $\mathrm{SL}(3, q)$. Using the order of a and the list of possible subgroups [Mit, Har] proves the claim.

6. SL(n, q)

We now turn to the general case of Theorem B for the groups $\mathrm{PSL}(n, q)$, using a variation on the approach in Section 5.

Theorem 6.1. *Let* $n \geq 4$.

(a) $\mathrm{SL}(n, q)$ *has a presentation with* 7 *generators*, 25 *relations and bit-length* $O(\log n + \log q)$.

(b) $\mathrm{PSL}(n, q)$ *has a presentation with* 7 *generators*, 26 *relations and bit-length* $O(\log n + \log q)$.

(c) $\mathrm{SL}(4, q)$ *has a presentation with* 6 *generators*, 21 *relations and bit-length* $O(\log q)$.

(d) $\mathrm{SL}(5, q)$ *has a presentation with* 6 *generators*, 22 *relations and bit-length* $O(\log q)$.

Proof. We use two presentations:

- The presentation $\langle X \mid R\rangle$ for $F = \mathrm{SL}(3, q)$ in Theorem 5.1. We view F as the group of matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ in $G = \mathrm{SL}(n, q)$ and only write the upper left 3×3 block.
- The presentation $\langle Y \mid S\rangle$ for $T = A_n$ in Theorem 3.36, where T acts on $\{1, \ldots, n\}$ (with X and Y disjoint). We view T as permutation matrices.

We will also use the subgroup $L = \mathrm{SL}(2,q)$ of F consisting of matrices in the upper left 2×2 block. We use the following elements:

$$c = \begin{pmatrix} 0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix},\ f = \begin{pmatrix} 0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix} \in F,$$

$a \in L$ such that $\langle a, a^f\rangle = L$,
$(1,2,3), (1,3)(2,4) \in T$, and
σ and $\tau = (1,2)(3,4)$ in T interchanging 1 and 2 and generating the set-stabilizer $T_{\{1,2\}}$ of $\{1,2\}$ in T.

Bit-length: c, f and a have bit-length $O(\log q)$ using Remark 4.6. We may assume that σ is a cycle of length $n-2$ or $n-3$ on $\{3,\dots,n\}$; both σ and τ can be viewed as words in Y of bit-length $O(\log n)$ (by Remark 3.33).

We will show that G is isomorphic to the group J having the following presentation.

Generators: X, Y.

Relations:

(1) R.
(2) S.
(3) $c = (1,2,3)$.
(4) $a^\sigma = a^f$.
(5) $a^\tau = a^f$.
(6) $(a^f)^\sigma = a$.
(7) $[a, a^{(1,3)(2,4)}] = 1$.
(8) $[a^f, a^{(1,3)(2,4)}] = 1$ (needed only when n is 4 or 5).

As usual, there is a surjection $\pi\colon J \to G$, and J has subgroups we will identify with $F = \langle X\rangle$ and $T = \langle Y\rangle$. Since τ has order 2, (5) implies that $(a^f)^\tau = a$. Hence, by (4)–(6), $\langle \sigma,\tau\rangle$ normalizes L, inducing the same automorphism group as $\langle f\rangle$ on L. In particular, elements of $\langle\sigma,\tau\rangle$ that fix 1 and 2 must centralize L, while elements interchanging 1 and 2 act as f.

It follows that $|L^T| \le \binom{n}{2}$; as usual, we use π to obtain equality. Then L^T can be identified with the set of all 2-sets of $\{1,\dots,n\}$. Its subset $L^{\langle c\rangle}$ consists of 3 subgroups corresponding to the 2-sets in $\{1,2,3\}$. Consequently, any two distinct members of L^T can be conjugated by a single element of T to one of the pairs $L, L^{(1,2,3)}$ or $L, L^{(1,3)(2,4)}$. Here $\langle L, L^{(1,2,3)}\rangle = \langle L, L^c\rangle = F$ by (3).

We will use (7)–(8) to show that $[L, L^{(1,3)(2,4)}] = 1$. By our comment about elements of $\langle\sigma,\tau\rangle$, we have $a^{(1,2)(5,6)} = a^f$ and $a^{(3,4)(5,6)} = a$. Then

$$\begin{aligned} 1 &= [a, a^{(1,3)(2,4)}]^{(1,2)(5,6)} &&= [a^f, (a^{(3,4)(5,6)})^{(1,3)(2,4)}] = [a^f, a^{(1,3)(2,4)}]\\ 1 &= [a, a^{(1,3)(2,4)}]^{(1,2)(3,4)} &&= [a^f, (a^f)^{(1,3)(2,4)}]\\ 1 &= [a^f, a^{(1,3)(2,4)}]^{(1,2)(3,4)} &&= [a, (a^f)^{(1,3)(2,4)}], \end{aligned} \tag{6.2}$$

where the first equations explain the comment in (8).

Thus, any two distinct members of L^T either generate a conjugate of $F = \mathrm{SL}(3,q)$ or commute. Consequently, $N := \langle L^T\rangle \cong G$ (see [GKKL1, Sections 5.1 or 5.2] for the Steinberg presentation). Moreover, $N \trianglelefteq J$, and J/N is a homomorphic image of $\langle Y\rangle \cong A_n$ in which c is sent to 1. Thus, $J/N = 1$.

The bit-length follows easily from those of $\langle X \mid R\rangle$ and $\langle Y \mid S\rangle$.

We still need to count the number of relations. If $n \geq 6$ then we have used $14+7+5$ relations by Theorems 3.36 and 5.1. However, *we can remove one generator and one relation as follows.* In Theorem 5.1 we used a generator "c" (corresponding to the permutation matrix acting as $(3,2,1)$). Hence we replace that element c by a word in Y representing $(1,2,3) \in T$ in the relations appearing in that theorem. Then (3) is implied by the *new* presentation $\langle X \mid R \rangle$.

If $n=4$ or 5 we use the presentation (3.2) with only 2 generators and 3 relations. Then the preceding presentation requires only $14+4+6-1$ relations. Moreover, when $n=4$ we can delete σ entirely, saving two further relations (4) and (6); and when $n=5$ we can choose σ of order 2 and delete (6).

Finally, we need to add one further relation in order to obtain $\mathrm{PSL}(n,q)$. Let $h_{i,j}$ be the matrix with ζ and ζ^{-1} in positions i and j, and 1 elsewhere. Let $m=(q-1)/(d,q-1)$. If n is odd, use Remark 3.10 to obtain the $(n-2)$-cycle $(2,\ldots,n-1)$, and then the additional relation $h_{1,n}^m\big(h_{2,n}^m(2,\ldots,n-1)\big)^{n-2}=1$ produces $\mathrm{PSL}(n,q)$ with the required bit-length. The case n even is similar. □

7. Remaining rank 2 groups

In this section we will provide presentations required in Theorem B for some of the rank 2 groups of Lie type. Since $\mathrm{P}\Omega^-(6,q) \cong \mathrm{PSU}(4,q)$, and we will handle all unitary groups in a different manner in Theorem 8.2, we only need to consider the groups $\mathrm{Sp}(4,q)$, $G_2(q)$, ${}^3D_4(q)$ and ${}^2F_4(q)$. Note that the last three groups do not appear inductively as Levi factors of any higher rank groups of Lie type.

7.1. $\mathbf{Sp(4,q)}$, $\boldsymbol{G_2(q)}$ and $\boldsymbol{{}^3D_4(q)}$.

Here the Weyl group is dihedral of order $2m = 8$ or 12.

Theorem 7.1. (a) $\mathrm{Sp}(4,q)$ *has a presentation with* 6 *generators*, 35 *relations and bit-length* $O(\log q)$.

(b) $\mathrm{PSp}(4,q) \cong \Omega(5,q)$ *has a presentation with* 6 *generators*, 36 *relations and bit-length* $O(\log q)$.

(c) $G_2(q)$ *and* ${}^3D_4(q)$ *have presentations with* 6 *generators*, 40 *relations and bit-length* $O(\log q)$.

Proof. (a) The root system Φ of $G=\mathrm{Sp}(4,q)$ has 8 roots, half of them long and half short. Let $\Pi=\{\alpha_1,\alpha_2\}$ be a set of fundamental roots with α_1 long; there are corresponding rank 1 groups $L_{\alpha_i} \cong \mathrm{SL}(2,q)$.

We assume that $q>9$ until the end of this proof, and use the presentation $\langle X_i \mid R_i \rangle$ of L_{α_i} in Theorem 4.5, with

$$X_i = \{u_{\alpha_i}, r_i, h_i\},\ i=1,2 \tag{7.2}$$

(here we are using r_i instead of t_i in order to approximate standard Lie notation; we assume that X_1 and X_2 are disjoint). The action of h_i on u_{α_i} is given in R_i, and $U_{\alpha_i} := \langle u_{\alpha_i}^{\langle h_i \rangle} \rangle$ has order q. Let $U_{\alpha_i}^\sharp := U_{\alpha_i} \setminus \{1\}$. There are $(2,q-1)$ orbits of $\langle h_i \rangle$ on $U_{\alpha_i}^\sharp$, with orbit representatives $u_{\alpha_i,1} := u_{\alpha_i}$ and $u_{\alpha_i,2}$ (where $u_{\alpha_i,2} := u_{\alpha_i}$ if q is even). If q is odd then $\langle h_1,h_2 \rangle$ has 2 orbits on both $U_{\alpha_1}^\sharp \times U_{\alpha_2}^\sharp$ and $U_{\alpha_1+\alpha_2}^\sharp \times U_{\alpha_2}^\sharp$, with respective orbit representatives $(u_{\alpha_1,a}, u_{\alpha_2})$, $a=1,2$, and $(u_{\alpha_1+\alpha_2,a}, u_{\alpha_2,a})$, $a=1,2$ (see [GKKL1, Lemma 5.3]).

The root groups $U_\alpha, \alpha \in \Phi$, will be built into our presentation.

We will show that G is isomorphic to the group J having the following presentation.

Generators: $X_1 \cup X_2$.

Relations:

(1) $R_1 \cup R_2$.

(2) $h_1^{r_2}$, $h_2^{r_1}$ and w^4 are explicit words in $\{h_1, h_2\}$, where $w := r_1 r_2$.

(3) $(u_{\alpha_i})^{h_j^{w^n}}$ is an explicit word in $u_{\alpha_i}^{\langle h_i \rangle}$ for $i \neq j$ and $0 \leq n < 4$.

Notation: $U_{\alpha_i} := \langle u_{\alpha_i}^{\langle h_i \rangle} \rangle$ for $i = 1, 2$, $H := \langle h_1, h_2 \rangle$, $N := \langle r_1, r_2, H \rangle$, and $W := N/H$ (which is isomorphic to D_8).

Label the subgroups $(U_{\alpha_i})^{w^n}$, $0 \leq n < 4$, in the usual way as

$$U_\alpha,\ \alpha \in \Phi := \{\pm\alpha_1, \pm\alpha_1, \pm(\alpha_1 + a_2), \pm(\alpha_1 + 2a_2)\}$$

[Cart, p. 46], so that $\Phi = \alpha_1^{\{w^n | 0 \leq n < 4\}} \cup \alpha_2^{\{w^n | 0 \leq n < 4\}}$.

For $i = 1, 2$ let $u_{\alpha_i,1}$ be another name for u_{α_i}, and let $u_{\alpha_i,2}$ stand for an explicit word in $u_{\alpha_i}^{\langle h_i \rangle}$ corresponding to an element of G described above. If $\alpha = \alpha_i^{w^n}$ for (unique) $i \in \{1, 2\}$ and $0 \leq n < 4$, write $u_{\alpha,a} := (u_{\alpha_i,a})^{w^n}$ for $a = 1, 2$.

(4) (a) $[u_{\alpha_1}, u_{\alpha_1+2\alpha_2}] = [u_{\alpha_1}, u_{\alpha_1+\alpha_2}] = 1$.

(b) $[u_{\alpha_1,a}, u_{\alpha_2}]$ is an explicit word in $u_{\alpha_1}^{\langle h_1 \rangle w^3} \cup u_{\alpha_2}^{\langle h_2 \rangle w^3} \subset U_{\alpha_1+2\alpha_2} \cup U_{\alpha_1+\alpha_2}$ for $a = 1, 2$.

(c) $[u_{\alpha_2,a}, u_{\alpha_1+\alpha_2,a}]$ is an explicit word in $u_{\alpha_1}^{\langle h_1 \rangle w^3} \subset U_{\alpha_1+2\alpha_2}$ for $a = 1, 2$.

First of all, the explicit words mentioned in these relations are obtained in G, and have bit-length $O(\log q)$ in the generators by Remark 4.6. It follows that this presentation has bit-length $O(\log q)$, and that there is a surjection $\pi\colon J \to G$. Moreover, this presentation has $|X_1| + |X_2| = 6$ generators and at most $|R_1| + |R_2| + 3 + 8 + 6 = 35$ relations.

By (1) there is a subgroup $L_{\alpha_i} \cong \mathrm{SL}(2, q)$ of J we can identify with $\langle X_i \rangle$.

By (2), $H \trianglelefteq N$ and $W \cong D_8$, as stated in (3).

Using L_{α_i} we see that $U_{\alpha_i}^{r_i} = U_{-\alpha_i}$. By (3), H normalizes each subgroup $U_\alpha = \langle u_\alpha^H \rangle$. It follows that W acts on Φ in the natural manner: there are 8 root groups U_α, $\alpha \in \Phi$, labeled as in (3) and permuted by N.

Fix roots α and $\beta \neq \pm\alpha$. Then H acts by conjugation on the set of all Chevalley commutator relations [GLS, p. 47], which we write as

$$[y_\alpha, y_\beta] = \textstyle\prod_\gamma v_\gamma \quad \text{with} \quad y_\alpha \in U_\alpha^\sharp,\ y_\beta \in U_\beta^\sharp,\ v_\gamma \in U_\gamma, \tag{7.3}$$

where γ runs through all positive integral combinations of α and β. Clearly (4) provides us with instances of (7.3) corresponding to each of the four W-orbits on pairs $\{\alpha, \beta\}$, $\beta \neq \pm\alpha$.

If $[y_\alpha, y_\beta] = 1$ with $\{y_\alpha, y_\beta\} = \{u_{\alpha_1}, u_{\alpha_1+2\alpha_2}\}$ or $\{u_{\alpha_1}, u_{\alpha_1+\alpha_2}\}$ in (4a), then $[U_\alpha, U_\beta] = [\langle y_\alpha^H \rangle, \langle y_\beta^H \rangle] = 1$. (For, the Chevalley commutator relations for G [GLS, p. 47] show that the corresponding fundamental subgroups of G commute. In particular, already in $\pi(J) = G$ suitable elements of $h_1^N \cup h_2^N$ act independently on U_α and U_β.)

If $[y_\alpha, y_\beta] \neq 1$ for the pairs $\{y_\alpha, y_\beta\}$ in (4b,c) then, by [GKKL1, Lemma 5.3], we obtain all relations (7.3) by conjugating the ones in (4b,c) by elements of H.

At this point we have verified the Steinberg-Curtis-Tits relations mentioned in Section 2. Thus, J is a homomorphic image of the universal group of Lie type for G, and hence is G [GLS, pp. 312-313].

(b) When q is odd, we just need one further relation to kill the center of $\mathrm{Sp}(4, q)$.

(c) These groups are handled in a manner similar to what was done in (a), replacing the number 4 by 6 at suitable places in order to obtain the Weyl group D_{12}. We sketch this very briefly since these groups do not arise in higher rank settings.

Once again we have fundamental subgroups $L_{\alpha_1} \cong \mathrm{SL}(2, q)$ and $L_{\alpha_2} \cong \mathrm{SL}(2, q)$ or $\mathrm{SL}(2, q^3)$, where the latter occurs for short α_2 in ${}^3D_4(q)$. We use versions of (1)–(4), labeling the roots as in [Cart, p. 46]. However, in order to avoid any use of $(3, q-1)$ elements $u_{\alpha_1,a}$ we proceed as follows in (4). The long root groups U_α of G generate a subgroup $\mathrm{SL}(3, q)$. Relations (5) and (6) from our presentation in Theorem 5.1, with $c := w^2$, are instances of (7.3); hence these are the only relations of this sort needed for long α, β. By [GKKL1, Lemma 5.3], this is the only situation where we might have needed to use several elements $u_{\alpha_1,a}$.

This time $W = N/H$ has seven orbits on pairs $\{\alpha, \beta\}$, $\beta \neq \pm\alpha$. Counting, we see that there are only 40 relations. Note the significant savings due to not needing more than one element of any root group U_α.

This proves the theorem when $q > 9$.

It remains to make some straightforward remarks about the remaining cases. For each of these, if L_{α_i} is a central extension of $\mathrm{PSL}(2, q)$ then it has a presentation $\langle Y_i \mid S_i\rangle$ using 2 generators and at most 3 relations [Sun, CMY, CR3]. We do not have to be concerned about lengths of words in these bounded situations, so we can assume that each of the elements $u_{\alpha_i}, u_{\alpha_i,a}, r_i, h_i$ is replaced by a word in Y_i. Once this has been done, we can again write the relations (1)–(4) in our new generating set, and then our previous argument goes through without difficulty. □

7.2. $\boldsymbol{{}^2F_4(q)}$. Here $q = 2^{2e+1} > 2$. There is no root system in the classical sense, but there are 16 "root groups" U_i, $1 \le i \le 16$. There are rank 1 groups $L_1 = \mathrm{Sz}(q)$ and $L_2 = \mathrm{SL}(2, q)$, and we use the presentation $\langle X_i \mid R_i\rangle$ for L_i in Section 4.5 or Theorem 4.5. If $i = 2$ then (7.2) holds, and $U_2 := \langle u_2^{\langle h_2\rangle}\rangle$ is elementary abelian of order q. On the other hand, X_1 has size 7 and contains elements u_1, r_1, h_1 behaving essentially as before, except that this time $U_1 = \langle u_1^{\langle h_1\rangle}\rangle$ is nonabelian of order q^2 with $Z(U_1) = \Phi(U_1) = \langle (u_1^2)^{\langle h_1\rangle}\rangle$ (where $\Phi(U_1)$ denotes the Frattini subgroup of U_1).

We use the following presentation.

Generators: $X_1 \cup X_2$.

Relations:

(1) $R_1 \cup R_2$.

(2) $w^8 = 1$, where $w := r_1 r_2$.

(3) $h_1^{r_2}$ and $h_2^{r_1}$ are explicit words in $\{h_1, h_2\}$.

(4) $(u_{\alpha_i})^{h_j^{w^n}}$ is an explicit word in $u_{\alpha_i}^{\langle h_i\rangle}$ for $\{i, j\} = \{1, 2\}$ and $0 \le n < 8$.
Notation: Let $u_{i+n} := (u_i)^{w^n}$ for $i = 1, 2$ and $1 \le n < 8$.

(5) $[u_i, u_j]$ is an explicit product of words in $u_k^{\langle h_1, h_2\rangle}$, $i < k < j$, for the pairs (i, j) with $i = 1$ and $2 \le j \le 8$, or $i = 2$ and $j = 4$, 6 or 8.

In the presented group J there are again subgroups L_i we can identify with $\langle X_i\rangle$, $i = 1, 2$. Also, $\langle r_1, r_2\rangle$ is dihedral of order 16 by (2), and normalizes $H := \langle h_1, h_2\rangle$ by (3). It then follows from (4) that H normalizes each subgroup $U_{i+n} := (U_i)^{w^n}$ for $i = 1, 2$ and $1 \le n < 8$.

As in the case of the other rank 2 groups, the known actions of $h_1^{w^n}$ and $h_2^{w^n}$ in (4) allow us to deduce from (5) an additional relation analogous to (7.3) for each pair of nontrivial cosets of the form $y_i\Phi(U_i)$, $y_j\Phi(U_j)$, for i, j as in (5) and $y_i \in U_i$, $y_j \in U_j$ (compare [GKKL1, Lemma 5.3]). By using the elementary identity $[x, uv] = [x, v][x, u]^v$, we see that these conjugates of the relations (5) imply all analogues of (7.3) for these i, j. Finally, we obtain further relations analogous to (7.3) by conjugating by elements of $\langle w\rangle$. It is now easy to see that we have all relations required for a presentation of G ([Gri, p. 412], [BGKLP, p. 105] and [GLS, p. 48] give the 10 formulas implicitly contained in (5)).

There are 16 relations (4) and 10 relations (5), for a total of $7 + 3$ generators and $43 + 9 + 1 + 2 + 16 + 10 = 81$ relations. As noted in Section 4.5, these numbers can easily be decreased to $4 + 3$ and $31 + 9 + 1 + 2 + 16 + 10$.

8. Unitary groups

Since the commutator relations for the odd-dimensional unitary groups are especially complicated (see, for example, [GLS, Theorem 2.4.5(c)]), we will deal with these groups separately. In fact, when combined with Theorems 3.36 and 4.10, recent presentations in [BeS] allow us to use surprisingly few generators and relations (cf. Theorem 8.2).

8.1. **Phan style presentations.** We will outline the presentation of $G = \mathrm{SU}(n, q)$, $n \ge 4$, in [BeS], based on one in [Ph]. In [BeS], subgroups $U_1, U_2 \cong \mathrm{SU}(2, q)$ of $\mathrm{SU}(3, q)$ are called a *standard pair* if U_1 and U_2 are the respective stabilizers in $\mathrm{SU}(3, q)$ of perpendicular nonsingular vectors.

Using an orthonormal basis, it is easy to see that G has subgroups $U_i \cong \mathrm{SU}(2, q)$, $1 \le i \le n - 1$, and $U_{i,j}$, $1 \le i < j \le n - 1$, satisfying the following conditions.

(P1) If $|j - i| > 1$ then $U_{i,j}$ is a central product of U_i and U_j.

(P2) For $1 \le i < n - 1$, $U_{i,i+1} \cong \mathrm{SU}(3, q)$, and U_i, U_{i+1} is a standard pair in $U_{i,i+1}$.

(P3) $G = \langle U_{i,j} \mid 1 \le i < j \le n - 1\rangle$.

The presentation we will use is the following analogue of the Curtis-Steinberg-Tits presentation mentioned in Section 2.

Theorem 8.1. [Ph, BeS] *If* (P1)–(P3) *hold in a group G, then G is isomorphic to a factor group of* $\mathrm{SU}(n, q)$ *in each of the following situations.*

(a) $q > 3$ *and* $n \ge 4$.

(b) $q = 2$ *or* 3, $n \ge 5$ *and the following hold*:

  (1) $\langle U_{i,i+1}, U_{i+1,i+2}\rangle \cong \mathrm{SU}(4, q)$ *whenever* $1 \le i \le n - 3$; *and*

  (2) *if* $q = 2$ *then*

    (i) $[U_i, U_{j,j+1}] = 1$ *whenever* $1 \le i \le n - 1$, $1 \le j \le n - 2$ *and* $i \ne j - 1, j, j + 1, j + 2$; *and*

    (ii) $[U_{i,i+1}, U_{j,j+1}] = 1$ *whenever* $1 \le i \le n - 2$, $1 \le j \le n - 2$ *and* $i \ne j - 2, j - 1, j, j + 1, j + 2$.

In [BeS], it is remarked that (P1)–(P3) do not provide a presentation of $\mathrm{SU}(n, 2)$. It is also noted that a standard pair in $\mathrm{SU}(3, 2)$ does not generate that group.

8.2. **Some specific presentations.** When $q = 2$ or 3 we need to deal with some small groups. The computer package MAGMA [CP] contains the following presentation:

$$\begin{aligned}\mathrm{SU}(4,3) = \langle x, y \mid x^3 &= y^8 = [y^4, x] = (xy)^7 = [x, (xy^{-1})^7] \\ &= [y^2, xy^2xy^2xy^2] = xx^{xyx^{-1}y^{-1}xy} = (xyx^{-1}y^2)^8\rangle = 1\rangle.\end{aligned}$$

We also need other small cases [Bray]:

$$\begin{aligned}\mathrm{SU}(6,2) = \langle a, b \mid a^2 = b^7 &= (ab^3)^{11} = [a,b]^2 \\ &= [a,b^2]^3 = [a,b^3]^3 = (ab)^{33} = (abab^2ab^3ab^{-3})^2 = 1\rangle. \\ \mathrm{SU}(5,2) = \langle a, b \mid a^2 = b^5 &= (ab)^{11} = [a,b]^3 = [a,b^2]^3 = [a,bab]^3 = [a,bab^2]^3 = 1\rangle. \\ \mathrm{SU}(4,2) = \langle a, b \mid a^2 = b^5 &= (ab)^9 = [a,b]^3 = [a,bab]^2 = 1\rangle.\end{aligned}$$

The last of these is [CMY, (10.8)].

8.3. **Presentations of unitary groups.** In this section we will prove the following

Theorem 8.2. *Let $n \geq 4$.*

(a) $\mathrm{SU}(n,q)$ *has a presentation with* 6 *generators*, 39 *relations and bit-length* $O(\log n + \log q)$.

(b) $\mathrm{PSU}(n,q)$ *has a presentation with* 6 *generators*, 40 *relations and bit-length* $O(\log n + \log q)$.

(c) $\mathrm{SU}(4,q)$ *and* $\mathrm{SU}(5,q)$ *have presentations with* 6 *generators*, 35 *relations and bit-length* $O(\log q)$.

(d) $\mathrm{PSU}(4,q)$ *and* $\Omega^-(6,q)$ *have presentations with* 6 *generators*, 36 *relations and bit-length* $O(\log q)$.

Proof. Let

- $F := \mathrm{SU}(m,q)$, with $m = 3, 4$ or 6 and $m \leq n$, and
- $A := A_n$, acting on $\{1, \dots, n\}$.

We view both of these groups as lying in $G = \mathrm{SU}(n,q)$, using an orthonormal basis of the underlying vector space: F consists of the matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ with an $m \times m$ block in the upper left corner, and A consists of permutation matrices.

For each m we assume that we have the following additional information:

- The presentation $\langle X \mid R\rangle$ of F in Theorem 4.10, except in the case of the pairs $(m,q) = (4,2), (4,3)$ or $(6,2)$, in which case $\langle X \mid R\rangle$ is given in Section 8.2.
- The presentation $\langle Y \mid S\rangle$ of A in Theorem 3.36 (where X and Y are disjoint).
- Two generators g, h for $W := \mathrm{SU}(2,q) < F$, where W consists of the matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ with a 2×2 block in the upper left corner, and where g and h can be viewed as words in X of bit-length $O(\log q)$ (using Remark 4.11).
- A word in X of bit-length $O(\log q)$ representing the element $c_{(1,2,3)} \in F$ that acts as the 3-cycle $(1,2,3)$ on the orthonormal basis (cf. Remark 4.11).
- A word in Y of bit-length $O(\log n)$ representing $(1,2,3)$ (cf. Remark 3.33).
- Permutations $\sigma = (1,2)(3,4)$ and τ in A that interchange 1 and 2 and generate the set-stabilizer S_{n-2} of $\{1,2\}$ in A, where σ and τ can be viewed as words in Y of bit-length $O(\log n)$ (using Remark 3.33).

Parts (b) and (d) of the theorem are handled at the end of the proof.

Case $q > 3$: Here we let $m = 3$. We will show that G is isomorphic to the group J having the following presentation. In view of the preceding remarks, this presentation has the desired bit-length.

Generators: $X \cup Y$.

Relations:

(1) $R \cup S$.

(2) $c_{(1,2,3)} = (1,2,3)$.

(3) $g^\sigma = g^\tau = g^{\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)}$ and $h^\sigma = h^\tau = h^{\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)}$. (The matrix $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ is not in W if q is odd, but can be viewed as inducing an automorphism of W.)

(4) $[W, W^{(13)(24)}] = 1$.

Since there is a surjection $\pi\colon J \to G$, there are subgroups of J we can identify with $F = \langle X\rangle$ and $A = \langle Y\rangle$.

By (3), $|W^A| \le \binom{n}{2}$, and using π we see that W^A consists of $\binom{n}{2}$ subgroups we will call $W_{i,j} = W_{j,i}$, $1 \le i < j \le n$, where $W_{i,j} \le F$ for $1 \le i < j \le 3$. If i,j,k,l are distinct, then (4) and the 4-transitivity of A imply that $[W_{i,j}, W_{k,l}] = 1$ and $\langle W_{i,j}, W_{i,k}\rangle \cong \langle W_{1,2}, W_{1,3}\rangle = F$.

Let $U_i := W_{i,i+1}$ and $U_{i,j} := \langle U_i, U_j\rangle$. These subgroups satisfy (P1)–(P2) and hence, by Theorem 8.1, $N := \langle U_{i,j} \mid 1 \le i < j \le n-1\rangle$ is a homomorphic image of G.

We claim that $N \trianglelefteq G$. For, $W_{1,3} \le \langle W_{1,2}, W_{2,3}\rangle$. By the 3-transitivity of A, it follows that $W_{i,k} \le \langle W_{i,j}, W_{j,k}\rangle$ for all distinct i,j,k. By induction, if $i < j+1$ then

$$\begin{aligned} W_{i,j} &\le \langle W_{i,j-1}, W_{j-1,j}\rangle \\ &\le \langle W_{i,i+1}, \dots, W_{j-2,j-1}, W_{j-1,j}\rangle = \langle U_i, \dots, U_{j-2}, U_{j-1}\rangle \le N. \end{aligned}$$

Thus, $N = \langle F^A\rangle \trianglelefteq J$. By (2), J/N is a quotient of A_n in which $(1,2,3)$ is mapped to 1. Thus, $J/N = 1$.

Total: $|X| + |Y| = 3+3$ generators and $|R| + |S| + 9 = 23 + 7 + 9$ relations.

This proves (a) when $q > 3$. For (c), we use the presentation (3.2) with only 2 generators and 3 relations. Hence, the preceding presentation requires only 23+3+9 relations.

Case $q = 3$: This time we let $m = 4$ and use the presentation $\langle X \mid R\rangle$ for $\mathrm{SU}(4,3)$ given in Section 8.2. We assume that $n \ge 5$, and that we have

- Words in X representing the elements $c_{(1,2,3)}, c_{(2,3,4)} \in F$ that act as the indicated permutations on the orthonormal basis.

We will show that G is isomorphic to the group J having the following presentation.

Generators: $X \cup Y$.

Relations:

(1′) $R \cup S$.

(2′) $c_{(1,2,3)} = (1,2,3)$, $c_{(2,3,4)} = (2,3,4)$.

(3′) $g^\sigma = g^\tau = g^{\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)}$ and $h^\sigma = h^\tau = h^{\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)}$.

By (2′) and (3′), $g^\sigma = g^\tau \in W$ and $h^\sigma = h^\tau \in W$. As before, it follows that W^A consists of $\binom{n}{2}$ subgroups $W_{i,j} = W_{j,i}$, $1 \le i < j \le n$, where $W_{i,j} \le F$ for $1 \le i < j \le 4$.

The previous relation (4) follows from the corresponding relation in $F = \mathrm{SU}(4,q)$.

If i,j,k,l are distinct, then (4) and the 4-transitivity of A give $[W_{i,j}, W_{k,l}] = 1$, $\langle W_{i,j}, W_{i,k}\rangle \cong \langle W_{1,2}, W_{1,3}\rangle \cong \mathrm{SU}(3,q)$ and $\langle W_{i,j}, W_{i,k}, W_{i,l}\rangle \cong \langle W_{1,2}, W_{1,3}, W_{1,4}\rangle = F$.

Once again, the subgroups $U_i := W_{i,i+1}$ and $U_{i,j} := \langle U_i, U_j \rangle$ of J satisfy (P1)–(P2). They also behave as in Theorem 8.1(b1) since $\langle U_{i,i+1}, U_{i+1,i+2} \rangle \cong \langle U_{1,2}, U_{2,3} \rangle = F$.

Once again, the subgroup N generated by all $U_{i,j}$ is isomorphic to G. As before, N is normal in $J = \langle X, Y \rangle$ and hence is J.

Total: $|X| + |Y| = 2 + 3$ generators and $|R| + |S| + 6 = 8 + 7 + 6$ relations.

Case $q = 2$: This time we let $m = 6$. Using the presentations in Section 8.2 we may assume that $n \geq 7$, and that we have

- Generators g', h' for $V := \mathrm{SU}(3,2)$ as words in X;
- Words in X representing elements $c_{(1,2,3)}, c_{(2,3,4,5,6)} \in F$ that act as the indicated permutations on the orthonormal basis; and
- Two permutations $\sigma' = (1,2)(4,5)$ and τ' that generate the set-stabilizer of $\{1,2,3\}$ in A and can be viewed as words in Y of bit-length $O(\log n)$ (cf. Remark 3.33).

We will show that G is isomorphic to the group J having the following presentation.

Generators: $X \cup Y$.

Relations:

(1″) $R \cup S$.

(2″) $c_{(1,2,3)} = (1,2,3)$, $c_{(2,3,4,5,6)} = (2,3,4,5,6)$.

(3″) $g^\sigma = g^\tau = g^{\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)}$ and $g^\sigma = h^\tau = h^{\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)}$.

(4″) $g'^{\tau'} = g'^{c[\tau']}$ and $h'^{\tau'} = g'^{c[\tau']}$, where $c[\tau']$ denotes the automorphism of V induced by the permutation matrix for the restriction of τ' to the first 3 coordinates.

As before, it follows from (2″) and (3″) that W^A consists of $\binom{n}{2}$ subgroups $W_{i,j} = W_{j,i}$, $1 \leq i < j \leq n$, where $W_{i,j} < F$ for $1 \leq i < j \leq 6$. The previous relation (4) follows from the corresponding relation in $F = \mathrm{SU}(6,2)$.

If i,j,k,l are distinct, then $W_{i,j}$ commutes with $W_{k,l}$ by (4) and the 4-transitivity of A.

By (2″), $\sigma' \in \langle c_{(1,2,3)}, c_{(2,3,4,5,6)} \rangle < F$, so that $V^{\sigma'} = V$. As before, (4″) then implies that V^A consists of $\binom{n}{3}$ subgroups $V_{i,j,k}$ for distinct $i,j,k \in \{1,\dots,n\}$, where $V_{i,j,k} \leq F$ for $1 \leq i,j,k \leq 6$ and $W_{i,j} \leq V_{i,j,k}$.

In view of the transitivity properties of A, if i,j,k,l,r,s are distinct then $\langle V_{i,j,k}, V_{i,j,l} \rangle \cong \langle V_{1,2,3}, V_{1,2,4} \rangle \cong \mathrm{SU}(4,2)$ and $[V_{i,j,k}, V_{l,r,s}] = 1$.

Let $U_i := W_{i,i+1}$, $U_{i,i+1} := V_{i,i+1,i+2}$ and $U_{i,j} := \langle U_i, U_j \rangle$ iff $|i-j| > 1$.

The subgroups $U_{i,j}$ satisfy (P1)–(P2). They also behave as in Theorem 8.1(b1) since $\langle U_{i,i+1}, U_{i+1,i+2} \rangle \cong \langle U_{1,2}, U_{2,3} \rangle = \mathrm{SU}(4,2)$. The conditions in Theorem 8.1(b2) also hold since they hold for the subgroups $U_i, U_{i,i+1}, U_{j,j+1}$ that lie in F.

Hence, the subgroup N generated by all $U_{i,j}$ is isomorphic to G. As before, N is normal in $J = \langle X, Y \rangle$ and hence is J.

Total: $|X| + |Y| = 2 + 3$ generators and $|R| + |S| + 8 = 8 + 7 + 8$ relations.

This proves (a) and (c) for all q. For (b) and (d) we note that a presentation of each group in question is obtained as in the proof of Theorem 6.1 by adding one new relation of bit-length $O(\log n + \log q)$ to a presentation in (a) or (c). □

9. General case

We now complete the proofs of Theorems A and B.

Theorem 9.1. *All finite simple groups of Lie type and rank $n \geq 3$ over $\mathbb{F}_q$ have presentations with at most* 14 *generators*, 76 *relations and bit-length* $O(\log n + \log q)$.

Proof. We use a variation on the methods in [GKKL1, Section 6.2]. By Theorems 5.1, 6.1 and 8.2, we may assume that G is neither a special linear nor unitary group.

Until the end of the proof we assume that our simple group G is the simply connected group of the given type. As usual, $\Pi = \{\alpha_1, \dots, \alpha_n\}$ is the set of fundamental roots of G, and for each i there are root groups $U_{\pm\alpha_i}$.

Case 1: G is a classical group. Number Π as in [GKKL1, Section 6.2]: the subsystem $\{\alpha_1, \dots, \alpha_{n-1}\}$ is of type A_{n-1}, α_n is an end node root and is connected to only one root α_j in the Dynkin diagram (here $j = n-1$ except for type D_n, when $j = n-2$). Let

(9.2) $\quad G_1 = \langle U_{\pm\alpha_i} \mid 1 \leq i < n \rangle$, $G_2 = \langle U_{\pm\alpha_n}, U_{\pm\alpha_j} \rangle$, $L_2 = \langle U_{\pm\alpha_n} \rangle$, $L = \langle U_{\pm\alpha_j} \rangle$.

Then G_1 has type A_{n-1} and G_2 is a rank 2 classical group.

Let L_1 be the subgroup of G_1 generated by the root groups that commute with L_2. Then L_1 is of type A_{n-2} unless G has type D_n, in which case L_1 is of type $A_1 \times A_{n-3}$.

We will use the following presentations (with X and Y disjoint):

- The presentation $\langle X \mid R \rangle$ for G_1 in Theorem 6.1; and
- The presentation $\langle Y \mid S \rangle$ for G_2 in Theorems 7.1 and 8.2(d). (The latter presentation is needed for the universal cover of $\Omega^-(6,q)$ rather than for general unitary groups.)

Choose two generators for L_1, two for L_2 and two for L, all viewed as words in X or Y.

Bit-length: Generators for the subgroup $\langle U_{\pm\alpha_1} \rangle \cong \mathrm{SL}(2,q)$ are used in the presentation $\langle X \mid R \rangle$ (cf. Section 6). Thus, by Remark 4.6, each element of $\langle U_{\pm\alpha_1} \rangle$ has bit-length $O(\log q)$ in X.

Recall that the presentation in Theorem 3.36 was used in the proof of Theorem 6.1. Hence, the cycles in Remark 3.33 have bit-length $O(\log n)$ in X. Since $L_2 = \langle U_{\pm\alpha_n} \rangle$ and $L = \langle U_{\pm\alpha_j} \rangle$ can each be obtained by conjugating $\langle U_{\pm\alpha_1} \rangle$ using one of those cycles, it follows that our generators for these groups have bit-length $O(\log n + \log q)$ in X. Similarly, in view of Remark 4.6 the proof of Theorem 7.1 shows that our two generators for L have bit-length $O(\log q)$ in terms of Y.

We need to be more careful with our choice of generators of L_1 in order to achieve the desired bit-length. We will assume that G does not have type D_n; the omitted case is very similar. If $n-1 = 2$, we can use any pair of generators by Remark 4.6. Assume that $n-1 \geq 3$. We temporarily write elements of $L_1 \cong \mathrm{SL}_{n-1}$ using $(n-1) \times (n-1)$ matrices. View $\langle U_{\pm\alpha_1} \rangle = \left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ with 2×2 blocks $*$, inside $\mathrm{SL}(3,q) = \langle U_{\pm\alpha_1}, U_{\pm\alpha_2} \rangle = \left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ with 3×3 blocks $*$. Choose $a \in \langle U_{\pm\alpha_1} \rangle$ such that $\mathrm{SL}(3,q) = \langle a, a^{(1,2,3)} \rangle$ (cf. Remark 5.7). The monomial transformation $g := (1,2,\dots,n-1)$ or $(1,2,\dots,n-1,-1,-2,\dots) = r_{12}\,(2,\dots,n-1)$ is in L_1; and g has bit-length $O(\log n + \log q)$ in X by the preceding paragraph. Note that $a^{(1,2,3)} = a^g$: since $n-1 \geq 3$, $(1,2,3)$ and g agree on the first two coordinates.

Then $L_1 \geq \langle a, g\rangle \geq \langle\langle a, a^g\rangle^{\langle g\rangle}\rangle = \langle\langle a, a^{(1,2,3)}\rangle^{\langle g\rangle}\rangle = \langle \mathrm{SL}(3,q)^{\langle g\rangle}\rangle = L_1$, where a and a^g have bit-length $O(\log n + \log q)$ in X.

We will show that G is isomorphic to the group J having the following presentation. In view of the preceding remarks, this presentation has the desired bit-length.

Generators: X, Y.

Relations:

(1) $R \cup S$.

(2) $[L_1, L_2] = 1$.
More precisely, impose the four obvious commutation relations using pairs of words in X or Y that map onto the chosen generators for L_1 or L_2.

(3) Identify the copies of L in G_1 and G_2.
More precisely, take the two generators for L, viewed as words in X and also in Y, and impose the equality of the corresponding words.

We claim that $J \cong G$. For, clearly J surjects onto G, and hence we may assume that $G_1 = \langle X\rangle$ and $G_2 = \langle Y\rangle$, L, L_1 and L_2 are subgroups of J. Clearly J is generated by the fundamental root groups contained in G_1 or G_2. Any two of these root groups satisfy the Curtis-Steinberg-Tits relations (see the references in Section 2): either they are both in G_1 or G_2, or they commute since $[L_1, L_2] = 1$. Thus, J is a homomorphic image of the universal finite group of Lie type of the given type, which proves the claim.

By Theorems 5.1, 6.1, 7.1 and 8.2(c), if the type is not D_n then the number of generators is $|X| + |Y| \leq 7 + 6$ and the number of relations is $|R| + |S| + 6 \leq 25 + 36 + 6 = 67$ (since 4 relations are required to ensure that $[L_1, L_2] = 1$ and 2 to identify the copies of L). For type D_n these numbers become $7+4$ and $25+14+6$.

As in [BGKLP, GKKL1], in each case we can kill the center of G with at most one additional relation of bit-length $O(\log n + \log q)$, except for some of the groups D_n, where two relations may be needed. Thus, in all cases we use at most 13 generators and at most 69 relations for all simple classical groups.

This proves the theorem for the classical groups. However, for use in $F_4(q)$, when $G \cong \mathrm{Sp}(6,q)$ we can use the presentation in Theorem 5.1 instead of the one in Theorem 6.1. This time $|X| + |Y| = 4 + 6$ and $|R| + |S| + 6 = 14 + 36 + 6 = 56$.

Case 2: G is an exceptional group. We modify the above argument slightly. If G is the universal cover of $E_n(q)$ with $6 \leq n \leq 8$, for a suitable numbering of Π we can define G_1, G_2, L_1, L_2, L essentially as in (9.2): G_1 still has type A_{n-1}, G_2 has type A_2, and this time L_1 has type $A_2 \times A_{n-4}$. Since L_1 is still generated by 2 elements, precisely as above we obtain $|X|+|Y| = 7+4$ and $|R|+|S|+6 = 25+14+6$.

If G is $F_4(q)$, then $G_1 = \mathrm{Sp}(6,q)$, and both G_2 and L_1 have type A_2. We just saw that G_1 has a presentation with 10 generators and 57 relations. By Theorem 5.1 we obtain $|X| + |Y| = 10 + 4$ and $|R| + |S| + 6 = 56 + 14 + 6 = 76$.

Similarly, if G is the universal cover of ${}^2E_6(q)$, then $G_1 = \mathrm{SU}(6,q)$ and G_2 and L_1 both have type A_2. Using Theorems 8.2(a) and 5.1, we obtain $|X| + |Y| = 8 + 4$ and $|R| + |S| + 6 = 43 + 14 + 6 = 63$.

Since $\langle X \mid R\rangle$ and $\langle Y \mid S\rangle$ have bit-length $O(\log q)$, the same is true of our presentation for G.

Once again we can kill the center of G with at most one additional relation of bit-length $O(\log q)$. $\square$

Proof of Theorems A and B. As pointed out in Section 1, Theorem A follows from Theorem B and [GKKL1, Lemma 2.1]. Theorem B is contained in Theorems 3.36, 4.5, 4.10, 5.1, 6.1, 8.2 and 9.1, together with Sections 4.5 and 7.2. □

10. Additional presentations of classical groups

We now provide an alternative to the approach in the preceding section for presentations of classical groups. We will use Section 3.6 and its notation concerning the group $W = W_n = \mathbb{Z}_2^{n-1} \rtimes A_n$, consisting of $n \times n$ real monomial matrices with respect to an orthonormal basis $v_1, \dots, v_n$ of $\mathbb{R}^n$.

10.1. **Groups of type $\boldsymbol{D_n}$.** By Theorem 6.1, since $\mathrm{P}\Omega^+(6,q) \cong \mathrm{PSL}(4,q)$ we only need to consider the case $n \geq 4$. We will use Proposition 3.42 to imitate the argument in Theorem 6.1.

Theorem 10.1. *The group $\Omega^+(2n,q)$ has a presentation with*
(a) 8 *generators*, 31 *relations and bit-length $O(\log n + \log q)$ if $n \geq 4$, and*
(b) 8 *generators*, 29 *relations and bit-length $O(\log q)$ if $n = 4$ or* 5.
At most one additional relation of bit-length $O(\log n + \log q)$ is needed in order to obtain a presentation for $\mathrm{P}\Omega^+(2n,q)$.

Proof. There is a hyperbolic basis $e_1, f_1, \dots, e_n, f_n$ of $V = \mathbb{F}_q^{2n}$ associated with $G = \Omega^+(2n,q)$. Then W consists of isometries. Moreover, W lies in G and permutes the pairs $\{e_i, f_i\}$, $1 \leq i \leq n$: if $n \geq 5$ then W is perfect, and for $n = 4$ we can see this by restricting from the group $\Omega^+(10,q)$.

Each element of W can be viewed using two different vector spaces: $\mathbb{R}^n$ and V. In the action on V, we write elements in terms of the standard orthonormal basis $v_1, \dots, v_n$. The resulting diagonal elements of W are the elements leaving each pair $\{e_i, f_i\}$ invariant. Since these two views are potentially confusing (especially when q is even), we will initially write elements in both manners.

We digress in order to observe that, *when q is odd, W does not lift to an isomorphic copy inside the universal cover $\hat{G}$ of G.* For, recall Steinberg's criterion [St2, Corollary 7.6]: an involution in $\Omega^+(2n,q)$ lifts to an element of order 4 in the spin group if and only if the dimension of its -1 eigenspace on V is $\equiv 2 \pmod 4$. Apply this to $\mathrm{diag}(-1,-1,1,\dots,1) = (e_1,f_1)(e_2,f_2) \in W$ in order to obtain an element of order 4 in $\hat{G}$, which proves our claim.

We view $F = \mathrm{SL}(3,q)$ as the subgroup of G preserving the subspaces $\langle e_1, e_2, e_3\rangle$ and $\langle f_1, f_2, f_3\rangle$ while fixing the remaining basis vectors. We use

- the presentation $\langle X \mid R\rangle$ for F in Theorem 5.1 and
- the presentation $\langle Y \mid S\rangle$ for $W = W_n$ in Proposition 3.42 (where X and Y are disjoint),

together with the following elements:

- $c = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in F$;
- $a \in F$ such that $L := \langle a, a^f\rangle \cong \mathrm{SL}(2,q)$ consists of all matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$ in F;
- $(3,2,1) = (e_3,e_2,e_1)(f_3,f_2,f_1)$, $(1,3)(2,4) = (e_1,e_3)(e_2,e_4)(f_1,f_3)(f_2,f_4) \in W$ and $s = \mathrm{diag}(-1,-1,1,\dots,1) = (e_1,f_1)(e_2,f_2) \in Y$; and

- σ and $\tau = (1,2)(3,4) = (e_1,e_2)(e_3,e_4)(f_1,f_2)(f_3,f_4) \in W$ that generate the subgroup of W fixing $\langle v_1 - v_2\rangle$ (within $\mathbb{R}^n$) and send $v_1 - v_2$ to $v_2 - v_1$.

Bit-length: c, f and a have bit-length $O(\log q)$ using Remark 4.6. We may assume that σ is the product of $\mathrm{diag}(-1,-1,-1,-1,1,\dots)$ and a cycle of odd length $n-2$ or $n-3$ on $\{3,\dots,n\}$; both σ and τ can then be viewed as words in Y of bit-length $O(\log n)$ (by Remark 3.33).

We will show that G is isomorphic to the group J having the following presentation.

Generators: X, Y.

Relations:

(1) R.
(2) S.
(3) $c = (3,2,1)$.
(4) $a^\sigma = a^f$, $a^\tau = a^f$.
(5) $(a^f)^\sigma = a$.
(6) $[a, a^{(1,3)(2,4)}] = 1$.
(7) $[a^f, a^{(1,3)(2,4)}] = 1$ if $n = 4$ or 5.
(8) $[a, a^{s^{(3,2,1)}}] = 1$.
(9) $[a^f, a^{s^{(3,2,1)}}] = 1$ if $n = 4$.

There is a surjection $J \to G$, in view of the previous remark concerning the universal cover of G together with the fact that the chosen element f acts on L in the same manner as any element of W that interchanges 1 and 2.

As usual, we may assume that $F = \langle X\rangle$ and $W = \langle Y\rangle$ lie in J. By (4) and (5), $\langle\sigma,\tau\rangle$ normalizes L (since τ has order 2 we also have $(a^f)^\tau = a$). As usual, it follows that L^W can be identified with the set of $n(n-1)$ pairs $\{\pm\alpha\}$ of vectors $\alpha = \pm v_i \pm v_j \in \mathbb{R}^n$, $i \neq j$, in the root system Φ of type D_n. The groups in L^W produce corresponding root groups X_α, $\alpha \in \Phi$.

Any unordered pair of distinct, non-opposite roots can be moved by W to one of the following pairs within $\mathbb{R}^n$:

(i) $v_1 - v_2$, $v_1 + v_2$ (ii) $v_1 - v_2$, $v_1 - v_3$
(iii) $v_1 - v_2$, $v_3 - v_1$ (iv) $v_1 - v_2$, $v_3 - v_4$.

Then the corresponding root groups can be moved in the same manner.

Let $N := \langle L^W\rangle = \langle X_\alpha \mid \alpha \in \Phi\rangle = \langle F^W\rangle \trianglelefteq J$.

We need to verify the Steinberg relations for the root groups X_α. The pairs (ii) and (iii) already lie in $F = \mathrm{SL}(3,q)$, so the desired relations are immediate. It remains to consider the cases (i) and (iv).

As in (6.2) (inside the proof of Theorem 6.1), (6) and (7) imply that the root groups determined by (iv) commute.

Before considering (i), we note that (4) and (5) imply that every element of W that interchanges $v_1 - v_2$ and $v_2 - v_1$ also interchanges a and a^f. Two such elements are s and, when $n \geq 5$, also $(1,2)(4,5)$. Since $t := s^{(3,2,1)} = \mathrm{diag}(1,-1,-1,1,\dots,1)$ commutes with s, and $t^{(1,2)(4,5)} = \mathrm{diag}(-1,1,-1,1,\dots,1) = st$, by (8) we have

$$1 = [a^s, (a^s)^t] = [a^f, (a^f)^t]$$
$$1 = [a^{(1,2)(4,5)}, a^{t(1,2)(4,5)}] = [a^f, (a^{(1,2)(4,5)s})^t] = [a^f, a^t].$$

By (9), the second of these relations also holds when $n = 4$. Then also $1 = [a^{ft}, a]$. Now the root groups $X_{v_1-v_2} < L_{v_1-v_2} = L = \langle a, a^f\rangle$ and $X_{v_1+v_2} < L_{v_1+v_2} = L^t = \langle a^t, a^{ft}\rangle$ commute, as required for (i).

Thus, $N \cong G$. Relation (3) pulls $(1, 2, 3)$ into N, so that $J/N = 1$.

Now use Proposition 3.42 in order to obtain a presentation for $\Omega^+(2n, q)$ having the stated numbers of generators and relations. Finally, at most one further relation of bit-length $O(\log n + \log q)$ is needed to kill the center. □

This should be compared to the presentations for these groups in Section 9 using 11 generators and 46 relations.

10.2. **Groups of type B_n and C_n.** This time we will glue W_n and a group of type B_2 or C_2.

Theorem 10.2. *The groups* $\mathrm{Sp}(2n, q)$, $\Omega(2n + 1, q)$ *and* $\Omega^-(2n + 2, q)$ *have presentations with*
(a) 10 *generators*, 58 *relations and bit-length* $O(\log n + \log q)$ *if* $n \geq 4$,
(b) 9 *generators*, 57 *relations and bit-length* $O(\log q)$ *if* $n = 4$ *or* 5, *and*
(c) 8 *generators*, 52 *relations and bit-length* $O(\log q)$ *if* $n = 3$.
At most one additional relation of bit-length $O(\log n + \log q)$ *is needed in order to obtain a presentation for* $\mathrm{PSp}(2n, q)$ *or* $\mathrm{P}\Omega^-(2n + 2, q)$.

Proof. The root system Φ of type C_n or B_n for $G = \mathrm{Sp}(2n, q)$, $\Omega(2n + 1, q)$ or $\Omega^-(2n + 2, q)$ consists of the vectors $\pm v_i \pm v_j$ for $1 \leq i < j \leq n$, and all $\pm 2v_i$ or $\pm v_i$, respectively. We may assume that a fundamental system is

$$\Pi = \{\alpha_1,\ \alpha_j = v_{j+1} - v_j \mid 2 \leq j \leq n - 1\},$$

where $\alpha_1 = 2v_1$ or v_1.

We will use the subgroup $L_{12} \cong \mathrm{Sp}(4, q)$, $\Omega(5, q)$ or $\Omega^-(6, q)$ corresponding to the root subsystem Φ_{12} generated by α_1 and α_2, and its rank 1 subgroups L_1 and L_2 determined by $\pm\alpha_1$ and $\pm\alpha_2$, respectively. We will also need the subgroup L_{23} corresponding to the root subsystem generated by α_2 and α_3.

There is a hyperbolic basis $e_1, f_1, \dots, e_n, f_n$ associated with G (with additional basis vectors v, or v and v', perpendicular to all of these in the orthogonal cases). We may assume that $W = W_n$ permutes these $2n$ vectors, with its normal subgroup $\mathbb{Z}_2^{n-1}$ fixing each pair $\{e_i, f_i\}$. We may assume that the support of L_{12} is $\langle e_1, f_1, e_2, f_2\rangle$ (or $\langle e_1, f_1, e_2, f_2, v\rangle$ or $\langle e_1, f_1, e_2, f_2, v, v'\rangle$ in the orthogonal cases); and if $n \geq 4$ then the support of $L_2^{(1,3)(2,4)}$ is $\langle e_3, e_4, f_3, f_4\rangle$, so that

$$[L_{12}, L_2^{(1,3)(2,4)}] = 1. \tag{10.3}$$

Similarly,

$$[L_1, L_2^{c^2}] = 1. \tag{10.4}$$

We will use the following elements, writing matrices for L_2 using the vectors e_1, e_2 (and, implicitly, also their "duals" f_1, f_2), and writing elements of W using $\mathbb{R}^n$:

- $c = (1, 2, 3) = (e_1, e_2, e_3)(f_1, f_2, f_3) \in W$ and $s = \mathrm{diag}(-1, -1, 1, \dots, 1) = (e_1, f_1)(e_2, f_2) \in Y$;
- $(2, 3, 4) = (e_2, e_3, e_4)(f_2, f_3, f_4)$, $(1, 3)(2, 4) = (e_1, e_3)(e_2, e_4)(f_1, f_3)(f_2, f_4) \in W$ if $n \geq 4$;

- $\sigma,\tau\in W$ generating the stabilizer $W_{\{\pm v_1,\pm v_2\}}=W_{\{\{e_1,f_1\},\{e_2,f_2\}\}}\cong \mathbb{Z}_2^{n-1}\rtimes S_{n-2}$;
- $h=\mathrm{diag}(\zeta^{-1},\zeta)$ generating the torus that normalizes the root subgroups $X_{\pm\alpha_2}$ of $L_2:=L_{\alpha_2}$;
- $u=\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\in X_{\alpha_2}$;
- $a\in L_2$ such that $L_2=\langle a,a^s\rangle$;
- $a_2\in L_2$ such that $L_2=\langle a_2,a_2^{r_2h}\rangle$;
- $b\in L_{12}$ such that $L_{12}=\langle b,b^s\rangle$; and
- two generators for L_1.

Bit-length: Once again, by Remarks 4.6 and 3.33 the stated elements of L_1, L_2 and L_{12} have bit-length $O(\log q)$, while σ and τ have bit-length $O(\log n)$.

The required elements a,a_2 and b exist. For b, use the fact that $\mathrm{PSp}(4,q)\cong \mathrm{P}\Omega(5,q)$ and then, in the orthogonal 5- or 6-dimensional group L_{12}, choose an element b of order $(q^2+1)/(2,q-1)$ such that $\langle b,b^s\rangle$ has no proper invariant subspace of dimension >1. (This means that $\langle b,b^s\rangle$ is irreducible, except in the case $\mathrm{P}\Omega(5,q)$, q even.)

We will use the following presentations:

- a presentation $\langle X\mid R\rangle$ for $L_{12}\cong \mathrm{Sp}(4,q)$, $\Omega(5,q)$ or $\Omega^-(6,q)$ when $G=\mathrm{Sp}(2n,q)$, $\Omega(2n+1,q)$ or $\Omega^-(2n+2,q)$, respectively; and
- a presentation $\langle Y\mid S\rangle$ for $W=W_n$ (where X and Y are disjoint).

We have corresponding root groups X_α whenever $\alpha\in\Phi_{12}$.

As in Section 10.1, in the orthogonal cases with q odd W does not lift to an isomorphic subgroup of the universal cover.

We will show that G is isomorphic to the group J having the following presentation.

Generators: X,Y.

Relations:

(1) R.
(2) S.
(3) x'^{σ} and x'^{τ} written as words in X, for each $x'\in\{b,b^s\}$.
(4) $(2,3,4)$ commutes with L_1 if $n\ge 4$.
(5) $c^{r_2}=r_2^2c^2$.
(6) $hh^ch^{c^2}=1$.
(7) $a_2^{h^c}=a_2^{\mathrm{diag}(1,\zeta^{-1})}$ written as a word in X.
(8) $[u^c,u]=(u^{r_2})^{c^2}$.
(9) $[u^c,u^{r_2}]=1$.
(10) $[b,a^{(1,3)(2,4)}]=1$ if $n\ge 4$.
(11) $[L_1,L_2^{c^2}]=1$ if $n=3$.
(12) s written as a word in $L_1\cup L_2$ if $n=3$.

Note that the relations (2)–(6) in Theorem 5.1 are the present relations (5)–(9) for $L_{23}:=\langle L_2,c\rangle\cong \mathrm{SL}(3,q)$. Also, the present relations (10) and (11) follow from (10.3) and (10.4). Therefore, as usual, there is a surjection $J\to G$, and we may assume that $L_{12}=\langle X\rangle$, $W=\langle Y\rangle$ and $L_{23}=\langle L_2,c\rangle$ lie in J.

Case $n\ge 4$**:** Relations (3) and (4) imply that $\langle W_{\{\pm v_1,\pm v_2\}},(2,3,4)\rangle = W_{\{\pm v_1\}}$ normalizes L_1. As usual, it follows that $|L_1^W|=n$. Similarly, by (3), $|L_{12}^W|=n(n-1)/2$ and each element of $W_{\{\pm v_1,\pm v_2\}}$ acts correctly on L_{12}. Then the normalizer of

L_2 in $W_{\{\pm v_1,\pm v_2\}}$ has index 2, so that $|L_2^W| \le n(n-1)$. As usual, it follows that $|L_2^W| = n(n-1)$. Starting with the root subgroups X_{α_1} and X_{α_2} of L_{12}, in J we obtain the "correct" set $X_{\alpha_1}^W \cup X_{\alpha_2}^W = \{X_\alpha \mid \alpha \in \Phi\}$ of root subgroups.

We will verify that the Steinberg relations hold for $N := \langle X_{\alpha_1}^W \cup X_{\alpha_2}^W \rangle = \langle L_{12}^W \rangle \trianglelefteq J$, after which we will have $J/N = 1$ by (10). Many of the required relations are already available for L_{12} and L_{23}.

Any unordered pair α, β of distinct, non-opposite roots can be moved by W to one of

(i) $\alpha_1, \pm\alpha_2$, (ii) $\alpha_1, \pm\alpha_1 \pm \alpha_2$, (iii) α_1, α_4, (iv) α_2, α_3, or (v) α_2, α_4.

Only pairs (iii) and (v) are not inside L_{12} or L_{23}.

Since each element of $W_{\{\pm v_1,\pm v_2\}}$ acts correctly on L_{12}, $s' := s^{(1,3)(2,4)}$ *commutes with* L_{12} (cf. (10.3)). By (10),

$$\begin{aligned}
1 = [b, a^{(1,3)(2,4)}]^s &= [b^s, a^{s'(1,3)(2,4)}] = [b^s, a^{(1,3)(2,4)}]\\
1 = [b, a^{(1,3)(2,4)}]^{s'} &= [b, (a^s)^{(1,3)(2,4)}]\\
1 = [b, a^{(1,3)(2,4)}]^{ss'} &= [b^{s's}, (a^{s's})^{(1,3)(2,4)}] = [b^s, (a^s)^{(1,3)(2,4)}].
\end{aligned}$$

Now $[L_{12}, L_2^{(1,3)(2,4)}] = [\langle b, b^s\rangle, \langle a, a^s\rangle^{(1,3)(2,4)}] = 1$, which takes care of the pairs (iii) and (v).

Thus, J is a central extension of G. We already noted that J cannot be the universal cover in the orthogonal cases. Hence, $J = N \cong G$.

For the counts in (a) and (b) use Theorems 7.1 and 8.2 together with Proposition 3.42.

Case $n = 3$: Once again, $|L_1^W| = 3$ and $|L_2^W| = 6$, and we obtain root groups X_α, $\alpha \in \Phi$.

This time we consider the subgroup $M := \langle L_{12}, L_{23}\rangle$ of $N := \langle X_{\alpha_1}^W \cup X_{\alpha_2}^W \rangle = \langle L_{12}^W \rangle \trianglelefteq J$. By the Curtis-Tits-Steinberg presentation mentioned in Section 2, in order to prove that $G \cong M$ we only need to consider pairs X_α, X_β of root groups with α, β lying in the subsystem of Φ determined by one of the pairs $\{\alpha_1, \alpha_2\}$, $\{\alpha_2, \alpha_3\}$ or $\{\alpha_1, \alpha_3\}$. In the first two cases the desired relations already hold in L_{12} or L_{23}. The last case is covered by (11).

It follows that $M \cong G$. As in (5.3), from (12) we obtain $W = \langle s, c\rangle \le M$. Then $J = N = M \cong G$. This time we use the presentation for $W_3 \cong A_4$ in (3.2). Hence, in (c) this presentation for J uses only $6 + 2$ generators and at most $36 + 3 + 7 + 4$ relations. □

Recall that we already had presentations for these groups in Section 9 having at most 13 generators and at most 67 relations.

11. Concluding remarks

1. Short and bounded presentations are goals of one aspect of Computational Group Theory ([Sim, pp. 290-291] and [HEO, p. 184]). Such presentations have various applications, such as in [LG, KS] for gluing together presentations in a normal series in order to obtain a presentation for a given matrix group. The presentations in the present paper are not short in the sense of length used in [Sim, HEO, GKKL1]. However, they have the potential advantage that they are simpler than those in [GKKL1], at least in the sense of requiring fewer relations. We hope that both types of presentations will turn out to be useful in Computational Group Theory.

2. The presentations for S_n and A_n in Section 3 that are related to prime numbers appear to be practical. The ones in Section 3.4 for general n have one unusual and awkward relation $y = w$, expressing y as a word in $X \cup X^y$; see (3.38) and the description of this word in the proof of Theorem 3.27. Experimentation appears to be needed in order to find a "nice" additional relation of this sort. That is, the presentation in Section 3.5 is among the easiest to describe of the presentations obtained using our methods, but it may not be the best in practice.

3. We used the presentations (3.2) of A_4 and A_5 in the proofs of Proposition 3.42 and Theorems 6.1, 8.2 and 10.2. If we had instead used the corresponding presentations of $\mathrm{SL}(2,3)$ or $\mathrm{SL}(2,5)$, with 2 generators and 2 relations, we could have saved an additional relation.

There are further small-rank cases where we could have proceeded in the same manner. Presentations are known for the universal central extension of A_n, $n \le 9$, with 2 generators and 2 relations [CHRR1, CHRR2]; and for A_{10} using 2 generators and 3 relations [Hav] (cf. Example 3.19(10)). For these small n such presentations can be used in our presentations in order to save several relations.

4. As in the proof of Corollary 3.40, Lemma 3.39 can be used in order to *decrease the number of relations by one* in Theorem A. The easiest way to see this is to use the "3/2-generation" of all finite simple groups [GK], according to which any one of our generators a is a member of a generating pair $\{a, b\}$; and then proceed as in Corollary 3.40. However, this uses an unnecessary amount of machinery, since each of our generating sets contains a member a for which a suitable b can be found without much difficulty.

5. For the groups S_n, A_n and $\mathrm{SL}(n,q)$ we were careful to make the number of relations small. For the other groups we were somewhat less careful. It is likely that one can obtain presentations of these groups with fewer relations using ideas provided here.

In particular, the presentations of rank 2 groups in Section 7 undoubtedly could be improved using the more careful approach in Section 5. For example, there should be no need for both $u_{\alpha_i,1}$ and $u_{\alpha_i,2}$.

The number of relations in Proposition 3.42 probably can be improved somewhat by using the ideas in Sections 3.1 and 3.4 in place of Theorem 3.36.

Phan-type presentations for orthogonal groups [BGHS, GHNS] should help decrease the numbers of relations in Sections 7, 9 and 10.

6. Theorem 10.1 did not deal with all central extensions of orthogonal groups, in particular, it did not handle spin groups. These can be dealt with in the same manner, by using a double cover of W_n.

7. There is a small generating subset of each finite simple group G producing a Cayley graph of diameter $O(\log |G|)$ [BKL, Ka]. Such generators need to be incorporated into Theorem A in order to obtain somewhat shorter presentations.

8. As observed in Section 3 we have constructed presentations for alternating and symmetric groups with bounded expo-length. The remaining presentations in this paper do not have this property, unless we only consider groups over *bounded* degree extensions of the prime field. An obstacle to our obtaining presentations with bounded expo-length is that we do not know sufficiently nice presentations of $\mathbb{F}_q$ when this field has large degree over the prime field $\mathbb{F}_p$.

In Section 4, we started with a presentation of $\mathbb{F}_q$ (as an algebra over $\mathbb{F}_p$) of the form $\mathbb{F}_q = \mathbb{F}_p[x,y]/\big(m(x), y - g_{\zeta^2}(x)\big)$, where x, y map onto ζ^{2d} and ζ^2, respectively, and used it to obtain a presentation of $\mathrm{SL}(2,q)$.

Roughly speaking, short (with length $O(\log q)$) presentations of $\mathbb{F}_q$ as an algebra over $\mathbb{F}_p$ yield presentations of $\mathrm{SL}(2,q)$ with short bit-length. However in order to obtain a presentation of $\mathrm{SL}(2,q)$ with bounded expo-length using the same method, we would need a presentation of $\mathbb{F}_q$ in which every relation involves only a bounded number of monomials. This is a computational question concerning "sparse" constructions of finite fields about which little appears to be known [GaN]. The same remarks also apply to the unitary and Suzuki groups.

Acknowledgements: We are grateful to George Havas for providing the presentations in Example 3.19(4, 10), to John Bray for providing the presentations in Section 8.2, and to Bob Griess for pointing out Steinberg's criterion [St2, Corollary 7.6] used in Section 10.1.

References

[Ar] E. Artin, Theorie der Zöpfe. Abh. Hamburg 4 (1925) 47–72.

[BGKLP] L. Babai, A. J. Goodman, W. M. Kantor, E. M. Luks and P. P. Pálfy, Short presentations for finite groups. J. Algebra 194 (1997) 79–112.

[BKL] L. Babai, W. M. Kantor and A. Lubotzky, Small diameter Cayley graphs for finite simple groups. European J. Combinatorics 10 (1989) 507–522.

[Bau] G. Baumslag, A finitely presented metabelian group with a free abelian derived group of infinite rank. Proc. Amer. Math. Soc. 35 (1972) 61–62.

[BGHS] C. Bennett, R. Gramlich, C. Hoffman and S. Shpectorov, Curtis-Phan-Tits theory, pp. 13–29 in: Groups, Combinatorics and Geometry (Durham, 1990; Eds. M. W. Liebeck and J. Saxl), Lond. Math. Soc. Lecture Note 165. Cambridge University Press, Cambridge 1992.

[BeS] C. D. Bennett and S. Shpectorov, A new proof of Phan's theorem. J. Group Theory 7 (2004) 287–310.

[Bn] D. J. Benson, Representations and cohomology. I. 2nd ed., Cambridge Stud. Adv. Math. 30, Cambridge University Press, Cambridge 1998.

[Bray] J. Bray, personal communication.

[BCLO] J. Bray, M. D. E. Conder, C. R. Leedham-Green and E. A. O'Brien, Short presentations for alternating and symmetric groups (preprint).

[Bre] R. Breusch, Zur Verallgemeinerung des Bertrandschen Postulates, dass zwischen x und $2x$ stets Primzahlen liegen. Math. Z. 34 (1932) 505–526.

[Bur] W. Burnside, Theory of Groups of Finite Order, 2nd ed. Cambridge University Press, Cambridge 1911.

[CR1] C. M. Campbell and E. F. Robertson, Classes of groups related to $F^{a,b,c}$. Proc. Roy. Soc. Edinburgh A78 (1977/78) 209–218.

[CR2] C. M. Campbell and E. F. Robertson, A deficiency zero presentation for $\mathrm{SL}(2,p)$. Bull. London Math. Soc. 12 (1980) 17–20.

[CR3] C. M. Campbell and E. F. Robertson, The efficiency of simple groups of order $< 10^5$. Comm. Algebra 10 (1982), 217–225.

[CRKMW] C. M. Campbell, E. F. Robertson, T. Kawamata, I. Miyamoto and P. D. Williams, Deficiency zero presentations for certain perfect groups. Proc. Roy. Soc. Edinburgh 103A (1986) 63–71.

[CRW] C. M. Campbell, E. F. Robertson and P. D. Williams, On presentations of $\mathrm{PSL}(2,p^n)$. J. Austral. Math. Soc. 48 (1990) 333–346.

[CHRR1] C. M. Campbell, G. Havas and C. Ramsay and E. F. Robertson, Nice efficient presentations for all small simple groups and their covers (to appear).

[CHRR2] C. M. Campbell, G. Havas and C. Ramsay and E. F. Robertson, On the efficiency of the simple groups with order less than a million and their covers (to appear).

[CMY] J. J. Cannon, J. McKay and K.-C. Young. The non-abelian simple groups G, $|G| < 10^5$ — presentations. Comm. Algebra 7 (1979) 1397–1406.

[CP] J. J. Cannon and C. Playoust, An introduction to Magma. School of Math. and Stat., Univ. Sydney, 1993.

[Carm] R. D. Carmichael, Introduction to the theory of groups of finite order. Ginn, Boston 1937.

[Cart] R. W. Carter, Simple groups of Lie type. Wiley, London 1972.

[CHR] M. Conder, G. Havas and C. Ramsay, Efficient presentations for the Mathieu simple group M_{22} and its cover, pp. 33–42 in: Finite Geometries, groups, and computation (Eds. A. Hulpke et al.), deGruyter, Berlin 2006.

[CoMo] H. S. M. Coxeter and W. O. J. Moser, Generators and relations for discrete groups, 3rd ed. Springer, New York-Heidelberg 1972.

[Cur] C. W. Curtis, Central extensions of groups of Lie type. J. reine angew. Math. 220 (1965) 174–185.

[Di] J. D. Dixon, The probability of generating the symmetric group. Math. Z. 110 (1969) 199–205.

[Er] P. Erdös, Über die Primzahlen gewisser arithmetischer Reihen. Math. Z. 39 (1935) 473–491.

[GaN] J. von zur Gathen and M. Nöcker, Polynomial and normal bases for finite fields. J. Cryptology 18 (2005) 337–355.

[GLS] D. Gorenstein, R. Lyons and R. Solomon, Amer. Math. Soc., Providence, R.I., 1998.

[Gr] R. Gramlich, Developments in finite Phan theory. Innov. Incidence Geom. (to appear).

[GHNS] R. Gramlich, C. Hoffman, W. Nickel and S. Shpectorov, Even-dimensional orthogonal groups as amalgams of unitary groups. J. Algebra 284 (2005) 141–173.

[Gri] R. L. Griess, Jr., Schur multipliers of finite simple groups of Lie type. Trans. Amer. Math. Soc. 183 (1973) 355–421.

[Gru] K. Gruenberg, Relation Modules of Finite Groups. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 25. Amer. Math. Soc., Providence, R.I., 1976.

[GrK] K. Gruenberg, and L. G. Koväcs, Proficient presentations and direct products of finite groups. Bull. Austral. Math. Soc. 60 (1999), 177–189.

[GH] R. M. Guralnick and C. Hoffman, The first cohomology group and generation of simple groups, Proceedings of a Conference on Groups and Geometry, Siena, Trends in Mathematics, 81-90, Birkäuser Verlag, 1998.

[GK] R. M. Guralnick and W. M. Kantor, Probabilistic generation of finite simple groups. J. Algebra 234 (2000) 743–792.

[GKKL1] R. M. Guralnick, W. M. Kantor, M. Kassabov and A. Lubotzky, Presentations of finite simple groups: a quantitative approach. J. Amer. Math. Soc. 21 (2008), 711–774. http://front.math.ucdavis.edu/0602.5508

[GKKL2] R. M. Guralnick, W. M. Kantor, M. Kassabov and A. Lubotzky, Presentations of finite simple groups: a cohomological and profinite approach. Groups, Geometry and Dynamics 1 (2007), 469–523 http://front.math.ucdavis.edu/0711.2817

[Har] R. W. Hartley, Determination of the ternary linear collineation groups whose coefficients lie in the GF(2^n). Ann. Math. 27 (1926) 140–158.

[Hav] G. Havas, personal communication.

[Ho] D. F. Holt, On the second cohomology group of a finite group. Proc. Lond. Math. Soc. 55 (1987) 22–36.

[HEO] D. F. Holt, B. Eick and E. A. O'Brien, Handbook of computational group theory. Chapman & Hall, Boca Raton 2005.

[HS] A. Hulpke and Á. Seress, Short presentations for three-dimensional unitary groups. J. Algebra 245 (2001) 719–729.

[Ka] W. M. Kantor, Some topics in asymptotic group theory, pp. 403-421 in: Groups, Combinatorics and Geometry (Durham, 1990; Eds. M. W. Liebeck and J. Saxl), Lond. Math. Soc. Lecture Note 165. Cambridge University Press, Cambridge 1992.

[KLu] W. M. Kantor and A. Lubotzky, The probability of generating a finite classical group. Geom. Dedicata 36 (1990) 67–87.

[KS] W. M. Kantor and Á. Seress, Computing with matrix groups, pp. 123–137 in: Groups, combinatorics and geometry (Durham, 2001; Eds. A. A. Ivanov et al.). World Sci. Publ., River Edge, NJ 2003.

[LG] C. R. Leedham-Green, The computational matrix group project, pp. 229–247 in: Groups and Computation III (Eds. W. M. Kantor and Á. Seress). deGruyter, Berlin-New York 2001.

[LiSh] M. W. Liebeck and A. Shalev, The probability of generating a finite simple group. Geom. Dedicata 56 (1995) 103–113.

[Lu] A. Lubotzky, Pro-finite presentations. J. Algebra 242 (2001), 672–690.

[Mil] G. A. Miller, Abstract definitions of all the substitution groups whose degrees do not exceed seven. Amer. J. Math. 33 (1911) 363–372.

[Mit] H. H. Mitchell, Determination of the ordinary and modular ternary linear groups. Trans. Amer. Math. Soc. 12 (1911) 207–242.

[Mol] K. Molsen, Zur Verallgemeinerung des Bertrandschen Postulates. Deutsche Math. 6 (1941) 248–256.

[Moo] E. H. Moore, Concerning the abstract groups of order $k!$ and $\frac{1}{2}k!$ holohedrically isomorphic with the symmetric and the alternating substitution groups on k letters. Proc. Lond. Math. Soc. 28 (1897) 357–366.

[Mor] P. Moree, Bertrand's postulate for primes in arithmetical progressions. Comput. Math. Appl. 26 (1993) 35–43.

[Neu] B. H. Neumann, On some finite groups with trivial multiplicator. Publ. Math. Debrecen 4 (1956) 190–194.

[Ph] K. W. Phan, On groups generated by three-dimensional special unitary groups I. J. Austral. Math. Soc. Ser. A23 (1977) 67–77.

[Ra] S. Ramanujan, A proof of Bertrand's Postulate. J. Indian Math. Soc. 11 (1919) 181–182.

[Ro] E. F. Robertson, Efficiency of finite simple groups and their covering groups, pp. 287–294 in: Finite groups–coming of age, Contemp. Math. 45, Amer. Math. Soc., Providence, R.I., 1985.

[Se] J.-P. Serre, Galois Cohomology, Springer, Berlin 2002.

[Sim] C. C. Sims, Computation with finitely presented groups. Cambridge Univ. Press, Cambridge 1994.

[Soi] L. Soicher (in preparation).

[St1] R. Steinberg, Lectures on Chevalley groups (mimeographed notes). Yale University 1967.

[St2] R. Steinberg. Generators, relations and coverings of algebraic groups, II. J. Algebra 71 (1981) 527–543.

[Sun] J. G. Sunday, Presentations of the groups $\mathrm{SL}(2,m)$ and $\mathrm{PSL}(2,m)$. Canad. J. Math. 24 (1972) 1129–1131.

[Suz] M. Suzuki, On a class of doubly transitive groups. Ann. of Math. 75 (1962) 105–145.

[Ti1] J. Tits, Les groupes de Lie exceptionnels et leur interprétation géométrique. Bull. Soc. Math. Belg. 8 (1956) 48–81.

[Ti2] J. Tits, Buildings of spherical type and finite BN-pairs. Springer, Berlin-New York 1974.

[To] J. A. Todd. A note on the linear fractional group. J. Lond. Math. Soc. 7, 195–200.

[Wi] J. S. Wilson, Finite axiomatization of finite soluble groups. J. Lond. Math. Soc. 74 (2006) 566–582.

Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532 USA
E-mail address: guralnic@usc.edu

Department of Mathematics, University of Oregon, Eugene, OR 97403 USA
E-mail address: kantor@math.uoregon.edu

Department of Mathematics, Cornell University, Ithaca, NY 14853-4201 USA
E-mail address: kassabov@math.cornell.edu

Department of Mathematics, Hebrew University, Givat Ram, Jerusalem 91904 Israel
E-mail address: alexlub@math.huji.ac.il